

; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1250434CD1
US-09-919-039-149

Query Match 82.9%; Score 34; DB 10; Length 827;
Best Local Similarity 87.5%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTPM 8
Db 562 MLAPVPM 569

RESULT 36
US-10-247-671-137
; Sequence 137, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 137
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1250434CD1
US-10-247-671-137

Query Match 82.9%; Score 34; DB 14; Length 827;
Best Local Similarity 87.5%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTPM 8
Db 562 MLAPVPM 569

RESULT 37
US-10-424-599-264106
; Sequence 264106, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 264106
; LENGTH: 57

; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80508C.1.pap
US-10-424-599-264106

Query Match 80.5%; Score 33; DB 15; Length 57;
Best Local Similarity 62.5%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTPM 8
Db 35 VIAPTEPM 42

RESULT 38
US-10-264-237-1843
; Sequence 1843, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patent in Ver. 3.1
; SEQ ID NO 1843
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-1843

Query Match 80.5%; Score 33; DB 15; Length 104;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTPM 8
Db 21 LLAPTEPM 28

RESULT 39
US-10-437-963-148789
; Sequence 148789, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 148789
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(147)
; OTHER INFORMATION: unsure at all Xaa locations

```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_49185C.1.pep
US-10-437-963-148789

Query Match      80.5%; Score 33; DB 16; Length 147;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
Db 128 ILAPTIFI 135

RESULT 40
US-10-437-963-201061
; Sequence 201061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 201061
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(168)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_96472C.1.pep
US-10-437-963-201061

Query Match      80.5%; Score 33; DB 16; Length 168;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
Db 144 ILAPTIFI 151

RESULT 41
US-10-437-963-109991
; Sequence 109991, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109991
```

```
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(169)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_14097C.1.pep
US-10-437-963-109991

Query Match      80.5%; Score 33; DB 16; Length 169;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
Db 139 ILAPTIFI 146

RESULT 42
US-10-437-963-109635
; Sequence 109635, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109635
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(173)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13774C.1.pep
US-10-437-963-109635

Query Match      80.5%; Score 33; DB 16; Length 173;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
Db 149 ILAPTIFI 156

RESULT 43
US-10-437-963-109941
; Sequence 109941, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

```
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109941
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(173)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_14050C.1.pep
US-10-437-963-109941

Query Match      80.5%; Score 33; DB 16; Length 173;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
Db 140 ILAPTIPI 147

RESULT 44
US-10-437-963-104575
; Sequence 104575, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104575
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(181)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101897C.1.pep
US-10-437-963-104575

Query Match      80.5%; Score 33; DB 16; Length 181;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
Db 139 ILAPTIPI 146

RESULT 45
US-10-437-963-109898
; Sequence 109898, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109898
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(195)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_14011C.1.pep
US-10-437-963-109898

Query Match      80.5%; Score 33; DB 16; Length 195;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
Db 138 ILAPTIPI 145

RESULT 46
US-10-437-963-202314
; Sequence 202314, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 202314
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_97605C.1.pep
US-10-437-963-202314

Query Match      80.5%; Score 33; DB 16; Length 201;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
Db 151 ILAPTIPI 158

RESULT 47
US-10-437-963-134864
; Sequence 134864, Application US/10437963
```

Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 134864
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(237)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_36596C.1.pep
US-10-437-963-134864

Query Match 80.5%; Score 33; DB 16; Length 237;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
:|||||
Db 19 ILVPTIPM 26

RESULT 48
US-10-437-963-137427
; Sequence 137427, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 137427
; LENGTH: 1557
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38912C.1.pep
US-10-437-963-137427

Query Match 80.5%; Score 33; DB 16; Length 1557;
Best Local Similarity 85.7%; Pred. No. 2.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIPTM 8
:|||||
Db 82 LAPTIPL 88

RESULT 49

US-10-437-963-204305
; Sequence 204305, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204305
; LENGTH: 2017
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2017)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99405C.1.pep
US-10-437-963-204305

Query Match 80.5%; Score 33; DB 16; Length 2017;
Best Local Similarity 75.0%; Pred. No. 2.7e+03;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
:|||||
Db 1280 ILAPTIPM 1287

RESULT 50
US-10-437-963-184153
; Sequence 184153, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 184153
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(139)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81173C.1.pep
US-10-437-963-184153

Query Match 78.0%; Score 32; DB 16; Length 139;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAPTIPM 8
| | | | |
Db 91 LVPTIPM 97

Search completed: February 9, 2005, 06:35:41
Job time : 112.649 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 05:55:37 ; Search time 44.4912 Seconds
(without alignments)
69.544 Million cell updates/sec

Title: US-10-032-361-4
Perfect score: 41
Sequence: 1 MLAPTIP 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 65 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	8	6	ABR82377
2	36	87.8	8	6	AAE30029
3	34	82.9	8	5	ABP54724
4	34	82.9	8	5	ABP54721
5	34	82.9	8	6	AAE30025
6	34	82.9	12	7	AAO23489
7	34	82.9	12	7	AAO23517
8	34	82.9	14	6	AAE30165
9	34	82.9	14	7	AAO23514
10	34	82.9	14	7	AAO23486
11	34	82.9	15	7	AAO23513
12	34	82.9	15	7	AAO23485
13	34	82.9	16	6	AAE30163
14	34	82.9	16	7	AAO23488
15	34	82.9	16	7	AAO23512
16	34	82.9	16	7	AAO23484
17	34	82.9	16	7	AAO23516
18	34	82.9	17	7	AAO23483
19	34	82.9	17	7	AAO23511
20	34	82.9	18	6	ABP57669
21	34	82.9	18	7	AAO23482
22	34	82.9	18	7	AAO23510
23	34	82.9	19	4	ABAB49912
24	34	82.9	19	6	AAE30167
25	34	82.9	19	6	AAE30144

26	34	82.9	19	6	AAE30162	Aae30162 Peptide #
27	34	82.9	19	6	AAE30172	Aae30172 Human HIF
28	34	82.9	19	6	AAE30158	Aae30158 HIF-1alpha
29	34	82.9	19	6	ABR82378	ABR82378 Hypoxia-i
30	34	82.9	19	8	ADP56728	ADP56728 Substrate
31	34	82.9	19	8	ADP79479	ADP79479 Hypoxia-i
32	34	82.9	20	6	ABP55440	ABP55440 Hypoxia-i
33	34	82.9	20	8	ADO22337	ADO22337 HIF-1alpha
34	34	82.9	28	7	AAO23532	AAO23532 Fluoresce
35	34	82.9	29	7	AAO23475	AAO23475 Murine HI
36	34	82.9	29	7	AAO23478	AAO23478 Murine HI
37	34	82.9	29	7	AAO23476	AAO23476 Murine HI
38	34	82.9	29	7	AAO23503	AAO23503 Murine HI
39	34	82.9	29	7	AAO23506	AAO23506 Murine HI
40	34	82.9	29	7	AAO23471	AAO23471 Murine HI
41	34	82.9	29	7	AAO23479	AAO23479 Murine HI
42	34	82.9	29	7	AAO23505	AAO23505 Murine HI
43	34	82.9	29	7	AAO23507	AAO23507 Murine HI
44	34	82.9	29	7	AAO23508	AAO23508 Murine HI
45	34	82.9	29	7	AAO23474	AAO23474 Murine HI
46	34	82.9	29	7	AAO23477	AAO23477 Murine HI
47	34	82.9	29	7	AAO23481	AAO23481 Murine HI
48	34	82.9	29	7	AAO23480	AAO23480 Murine HI
49	34	82.9	29	7	AAO23499	AAO23499 Murine HI
50	34	82.9	29	7	AAO23500	AAO23500 Murine HI
51	34	82.9	29	7	AAO23504	AAO23504 Murine HI
52	34	82.9	29	7	AAO23509	AAO23509 Murine HI
53	34	82.9	30	6	ABR82380	ABR82380 Hypoxia-i
54	34	82.9	34	4	ABAB49913	ABAB49913 Human/mur
55	34	82.9	34	6	AAE30161	AAE30161 Peptide #
56	34	82.9	34	6	AAE30151	AAE30151 HIF1alpha
57	34	82.9	34	3	AAV94637	AAV94637 HIF-1alpha
58	34	82.9	34	7	AAO23495	AAO23495 Murine HI
59	34	82.9	34	7	AAO23490	AAO23490 Murine HI
60	34	82.9	34	7	AAO23520	AAO23520 Murine HI
61	34	82.9	34	7	AAO23530	AAO23530 Murine HI
62	34	82.9	34	7	AAO23491	AAO23491 Murine HI
63	34	82.9	34	7	AAO23498	AAO23498 Murine HI
64	34	82.9	34	7	AAO23528	AAO23528 Murine HI
65	34	82.9	34	7	AAO23525	AAO23525 Murine HI

ALIGNMENTS

RESULT 1				
ABR82377				
ID	ABR82377	standard; peptide; 8 AA.		
XX	AC	ABR82377;		
XX	DT	06-NOV-2003	(first entry)	
XX	DE	Hypoxia-inducible factor 1 (HIF-1) alpha peptide inhibitor.		
XX	KW	HIF-1; hypoxia-inducible factor 1; HIF-1 alpha; ubiquitination; EPO;		
XX	KW	erythropoietin; vascular endothelial growth factor; VEGF; glycolytic;		
KW	KW	tranquillizer; vulnerary; cardiant; cerebroprotective; angiogenesis.		
XX	OS	Synthetic.		
XX	Key	Location/Qualifiers		
FT	Modified-site	4		
FT		/label= Hyp		
FT		/note= "hydroxyproline"		
XX	WO2003057820-A2.			
XX	17-JUL-2003.			
XX	04-OCT-2002; 2002WO-US031699.			
PF				
XX	21-DEC-2001; 2001US-00032361.			
XX	PR			

XX PA (KIMB) KIMBERLY-CLARK WORLDWIDE INC.
 XX PI Mcgrath K;
 XX XX WPI; 2003-645988/61.
 XX DR Novel peptide inhibitor of hypoxia-inducible factor 1 alpha
 XX PT ubiquitination, and activator of vascular endothelial growth factor
 XX PT transcription useful for treating tissue injuries including wounds,
 XX PT surgical incisions.
 XX PS Claim 3; Page 8; 37pp; English.
 XX CC The invention relates to peptide inhibitors of hypoxia-inducible factor
 CC (HIF-1) alpha ubiquitination. The peptide inhibitors thereby activate the
 CC transcription of erythropoietin (EPO), vascular endothelial growth factor
 CC (VEGF) and certain glycolytic enzymes. The peptide inhibitors are useful
 CC for treating tissue injuries including wounds, surgical incisions,
 CC chronic wounds, heart disease and stroke. The present sequence represents
 CC a specific example of HIF-1 alpha peptide inhibitor
 XX SQ Sequence 8 AA;
 Query Match 100.0%; Score 41; DB 6; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;
 QY 1 MLAPTIPM 8
 Db 1 MLAPTIPM 8
 RESULT 2
 ID AAE30029
 AC AAE30029 standard; peptide; 8 AA.
 XX DT 24-FEB-2003 (first entry)
 XX DE Human hypoxia-inducible factor 1 (HIF-1) alpha derived peptide #5.
 KW Entity localisation; light-generating fusion protein; LGP; diabetes;
 KW hypoxic tissue imaging; detection; cancer; hypoxia; heart disease;
 KW stroke; human; hypoxia-inducible factor 1; HIF-1 alpha.
 XX OS Homo sapiens.
 XX PN WO200275278-A2.
 XX PD 26-SEP-2002.
 XX PF 20-MAR-2002; 2002WO-US008964.
 XX PR 20-MAR-2001; 2001US-0277425P.
 XX PR 20-MAR-2001; 2001US-0277431P.
 XX PR 20-MAR-2001; 2001US-0277440P.
 XX PR 09-NOV-2001; 2001US-0332334P.
 XX PR 09-NOV-2001; 2001US-0332493P.
 XX PR 09-NOV-2001; 2001US-0345200P.
 XX PR 20-DEC-2001; 2001US-0342598P.
 XX PR 20-DEC-2001; 2001US-0345131P.
 XX PR 20-DEC-2001; 2001US-0345132P.
 XX PR 19-MAR-2002; 2002US-00101662.
 XX (DAND) DANA FARBER CANCER INST INC.
 XX PI Kaelin WG, Livingston DM, Kim W;
 XX WPI; 2003-018815/01.
 XX DR Detecting localization of an entity e.g. hypoxic tissue, tumor, wound in

PT a subject, by using a light-generating fusion protein having a ligand
 PT binding site and light-generating polypeptide moiety.
 XX Example 1; Fig 4B; 129pp; English.
 XX CC The invention relates to a method of detecting localisation of an entity
 CC in a subject. The method involves administering to the subject a light-
 CC generating fusion protein (LGP), or a cell expressing LGP, where LGP
 CC comprises a ligand binding site and a light-generating polypeptide
 CC moiety, and light generation of LGP changes upon binding of a ligand at
 CC the ligand binding site, allowing for co-localisation of LGP and an
 CC entity, and imaging localised LGP. The method is useful for detecting the
 CC localisation of an entity, such as a molecule, macromolecule, polymer,
 CC protein, antibody, protein complex, polysaccharide, nucleic acid,
 CC particle, inert material, organelle, cell, embryo, microorganism,
 CC bacteria, virus, fungus, prion, tumour, tissue, cellular environment
 CC comprising damaged tissue, diseased tissue or hypoxic tissue, wound,
 CC organ, proliferating cell and pathogen in a subject. It is particularly
 CC useful for measuring prolyl hydroxylase activity; for imaging hypoxic
 CC tissue and for detecting cancerous tissue. LGPs are useful for screening
 CC modulators of activity or latency of (or predisposition to) disorders
 CC such as hypoxia, cancer, diabetes, heart disease or stroke. The present
 CC sequence is human hypoxia-inducible factor 1 (HIF-1) alpha derived
 CC peptide. This sequence is used to illustrate the method of the invention
 XX SQ Sequence 8 AA;
 Query Match 87.8%; Score 36; DB 6; Length 8;
 Best Local Similarity 87.5%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 1;
 QY 1 MLAPTIPM 8
 Db 1 MLAPTIPM 8
 RESULT 3
 ID ABP54724 standard; peptide; 8 AA.
 XX AC ABP54724;
 XX DT 30-DEC-2002 (first entry)
 XX DE Hypoxia inducible factor 1-alpha peptide.
 XX KW Hypoxia inducible factor 1; HIF-1; human; transcription factor;
 KW vasotrophic; cardiatic; thrombolytic; gene therapy.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Modified-site 4 /label= OTHER
 FT /note= "hydroxyproline"
 XX PN WO200274980-A2.
 XX PD 26-SEP-2002.
 XX PF 20-MAR-2002; 2002WO-US008946.
 XX PR 20-MAR-2001; 2001US-0277425P.
 XX PR 20-MAR-2001; 2001US-0277431P.
 XX PR 20-MAR-2001; 2001US-0277440P.
 XX PR 09-NOV-2001; 2001US-0332334P.
 XX PR 09-NOV-2001; 2001US-0332493P.
 XX PR 09-NOV-2001; 2001US-0345200P.
 XX PR 20-DEC-2001; 2001US-0342598P.
 XX PR 20-DEC-2001; 2001US-0345131P.
 XX PR 20-DEC-2001; 2001US-0345132P.
 XX PR 19-MAR-2002; 2002US-00101816.

PT Kaelin WG, Livingston DM, Kim W;
XX WPI; 2003-018815/01.
DR
XX
PT Detecting localization of an entity e.g. hypoxic tissue, tumor, wound in
PT a subject, by using a light-generating fusion protein having a ligand
PT binding site and light-generating polypeptide moiety.
XX
XX Example 1; Fig 4B; 129pp; English.
PS
XX The invention relates to a method of detecting localisation of an entity
CC in a subject. The method involves administering to the subject a light-
CC generating fusion protein (LGP), or a cell expressing LGP, where LGP
CC comprises a ligand binding site and a light-generating polypeptide
CC moiety, and light generation of LGP changes upon binding of a ligand at
CC the ligand binding site, allowing for co-localisation of LGP and an
CC entity, and imaging localised LGP. The method is useful for detecting the
CC localisation of an entity, such as a molecule, macromolecule, polymer,
CC protein, antibody, protein complex, polysaccharide, nucleic acid,
CC particle, inert material, organelle, cell, embryo, microorganism,
CC bacteria, virus, fungus, prion, tumour, tissue, cellular environment
CC comprising damaged tissue, diseased tissue or hypoxic tissue, wound,
CC organ, proliferating cell and pathogen in a subject. It is particularly
CC useful for measuring prolyl hydroxylase activity; for imaging hypoxic
CC tissue and for detecting cancerous tissue. LGPs are useful for screening
CC modulators of activity or latency of (or predisposition to) disorders
CC such as hypoxia, cancer, diabetes, heart disease or stroke. The present
CC sequence is human hypoxia-inducible factor 1 (HIF-1) alpha derived
CC peptide. This sequence is used to illustrate the method of the invention
XX
SQ Sequence 8 AA;
Query Match 82.9%; Score 34; DB 6; Length 8;
Best Local Similarity 87.5%; Pred. No. 1.8e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAPTIPM 8
DB 1 MLAPYIPM 8
RESULT 6
AAO23489
ID AAO23489 standard; peptide; 12 AA.
XX
AC AAO23489;
XX
DT 12-FEB-2004 (first entry)
XX
DE Murine HIF-1alpha protein N-TAD region fragment (residues 560-571).
XX
XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW N-TAD.
XX
OS Mus sp.
XX
XX WO2003074560-A2.
XX
PD 12-SEP-2003.
XX
XX 05-MAR-2003; 2003WO-SE000372.
XX
XX 05-MAR-2002; 2002US-0361333P.
XX
XX 27-NOV-2002; 2002US-0429307P.
XX
XX (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
XX Pereira T, Poellinger L, Hellstroem M;
XX
XX WPI; 2003-712876/67.
XX
PT New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.
XX
XX Claim 39; Fig 22; 96pp; English.
XX
XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1alpha underexpression in a
CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell or a group of cells,
CC increasing angiogenetic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. Sequences AAO23482-489 represent fragments within
CC the N-TAD region of a murine HIF-1 alpha protein, that can be used for
CC treatment for hypoxic-related conditions
XX
SQ Sequence 12 AA;
Query Match 82.9%; Score 34; DB 7; Length 12;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAPTIPM 8
DB 1 MLAPYIPM 8
RESULT 7
AAO23517
ID AAO23517 standard; peptide; 12 AA.
XX
AC AAO23517;
XX
DT 12-FEB-2004 (first entry)
XX
DE Murine HIF-1alpha protein N-TAD region fragment (residues 560-571).
XX
XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW N-TAD.
XX
OS Mus sp.
XX
XX WO2003074560-A2.
XX
PD 12-SEP-2003.
XX
XX 05-MAR-2003; 2003WO-SE000372.
XX
XX 05-MAR-2002; 2002US-0361333P.
XX
XX 27-NOV-2002; 2002US-0429307P.
XX
XX (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
XX Pereira T, Poellinger L, Hellstroem M;
XX
XX WPI; 2003-712876/67.
XX
XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.
XX

PS Example 8; Fig 22; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1 α protein that has (a) an altered transactivation capacity and improved stability at normoxia. The HIF-1 α protein, polynucleotide, vector, and pharmaceutical composition are useful for increasing angiogenesis, interfering with a normal response to reoxygenation following hypoxia, or treating a condition associated with HIF-1 α underexpression in a cell, a group of cells, or an organism, e.g. ischaemia, diabetic retinopathy, inflammation, coronary heart disease, rheumatoid arthritis, stroke. The proteins and pharmaceutical compositions are also useful for mimicking the hypoxic response or artificially inducing a hypoxic response in a cell, group of cells, or organism, inducing vascular formation or vascular development in a cell or a group of cells, increasing angiogenic activity in a cell, or influencing erythropoietin production, metabolism, or an inflammatory response in a cell, a group of cells, or an organism. Sequences AAO23510-517 represent fragments within the N-TAD region of a murine HIF-1 α protein.

XX Sequence 12 AA;

Query Match 82.9%; Score 34; DB 7; Length 12;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
|||||
DB 1 MLAPYIPM 8

RESULT 8
AAE30165
ID AAE30165 standard; peptide; 14 AA.

AC AAE30165;

DT 24-FEB-2003 (first entry)

DE Peptide #5 used to block HIF-1 α /pVHL interaction.

KW Hypoxia inducible factor; HIF hydroxylase; inflammatory disorder; cancer;
wound healing; ischaemia; transplantation; blood pressure; gene therapy.

XX Unidentified.

OS WO200274981-A2.

PN 26-SEP-2002.

PD 21-MAR-2002; 2002WO-GB001381.

PP 21-MAR-2001; 2001GB-00007123.

PR 02-AUG-2001; 2001GB-00018952.

XX (ISIS-) ISIS INNOVATION LTD.

PA Maxwell PH, Pugh CW, Ratcliffe PJ, Schofield CJ;

PI WPI; 2003-018808/01.

XX Novel isolated polypeptide useful for treating ischemia, wound healing,
PT auto-, allo-, and xeno-transplantation, systemic high blood pressure,
PT cancer, or inflammatory disorders.

XX Example 1; Page 247; 256pp; English.

XX The invention relates to polypeptides having hypoxia inducible factor
CC (HIF) hydroxylase activity, referred to as PHD polypeptides (PHD 1,2 and
CC 3) and nucleic acid molecules encoding such polypeptides. Polypeptides of
CC the invention are used for treating conditions such as ischaemia, wound
CC healing, auto-, allo-, and xeno-transplantation, systemic high blood
CC pressure, cancer, or inflammatory disorders. They are useful in anti-
CC sense regulation of the HIF hydroxylase activity and in particular HIF

CC prolyl hydroxylase activity within a cell. They are also used to identify
CC additional substrates of HIF hydroxylases. Sequences of the invention are
CC used to design double stranded RNAs for use in RNA interference. They are
CC used as therapeutic agents and in purification, isolation, or screening
CC methods involving immuno-precipitation techniques and for detecting
CC polypeptides in biological samples. The invention is useful in gene
CC therapy. The present sequence is a peptide used to block HIF-1 α /pVHL
CC interaction. This sequence is used in the invention

XX Sequence 14 AA;

Query Match 82.9%; Score 34; DB 6; Length 14;

Best Local Similarity 87.5%; Pred. No. 23;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8

DB 6 MLAPYIPM 13

RESULT 9

AAO23514

ID AAO23514 standard; peptide; 14 AA.

AC AAO23514;

DT 12-FEB-2004 (first entry)

DE Murine HIF-1 α protein N-TAD region fragment (residues 560-573).

XX HIF-1 α ; hypoxia-inducible factor -1 α ; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW N-TAD.

XX Mus sp.

OS WO2003074560-A2.

PN 12-SEP-2003.

PD 05-MAR-2003; 2003WO-SE000372.

PF 05-MAR-2002; 2002US-0361333P.

PR 27-NOV-2002; 2002US-0429307P.

XX (ANGI-) ANGIOGENETICS SWEDEN AB.

XX Pereira T, Poellinger L, Hellstroem M;

PI WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1 α protein, useful for increasing
CC angiogenesis, or treating a condition associated with HIF-1 α
CC underexpression in a cell, a group of cells, or an organism, e.g.
CC ischemia or inflammation.

XX Example 8; Fig 22; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1 α protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1 α protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1 α underexpression in a
CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell or a group of cells,
CC increasing angiogenic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of

CC cells, or an organism. Sequences AAO23510-517 represent fragments within
 CC the N-TAD region of a murine HIF-1 alpha protein

XX SQ Sequence 14 AA;

Query Match 82.9%; Score 34; DB 7; Length 14;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
 |||||
 Db 1 MLAPVIPM 8

RESULT 10
 AAO23486
 ID AAO23486 standard; peptide; 14 AA.

XX AC AAO23486;

XX DT 12-FEB-2004 (first entry)

XX DE Murine HIF-1alpha protein N-TAD region fragment (residues 560-573).

XX KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW N-TAD.

XX OS Mus sp.

XX PN WO2003074560-A2.

XX PD 12-SEP-2003.

XX PF 05-MAR-2003; 2003WO-SE000372.

XX PR 05-MAR-2002; 2002US-0361333P.

XX PR 27-NOV-2002; 2002US-0429307P.

XX PA (ANGI-) ANGIOGENETICS SWEDEN AB.

XX PI Pereira T, Poellinger L, Hellstroem M;

XX DR WPI; 2003-712876/67.

XX PS New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.

XX PS Claim 39; Fig 22; 96pp; English.

XX CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. Sequences AAO23482-489 represent fragments within
 CC the N-TAD region of a murine HIF-1 alpha protein, that can be used for
 CC treatment for hypoxic-related conditions

XX SQ Sequence 14 AA;

Query Match 82.9%; Score 34; DB 7; Length 14;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
 |||||
 Db 1 MLAPVIPM 8

RESULT 11

AAO23513
 ID AAO23513 standard; peptide; 15 AA.

XX AC AAO23513;

XX DT 12-FEB-2004 (first entry)

XX DE Murine HIF-1alpha protein N-TAD region fragment (residues 559-573).

XX KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW N-TAD.

XX OS Mus sp.

XX PN WO2003074560-A2.

XX PD 12-SEP-2003.

XX PF 05-MAR-2003; 2003WO-SE000372.

XX PR 05-MAR-2002; 2002US-0361333P.

XX PR 27-NOV-2002; 2002US-0429307P.

XX PA (ANGI-) ANGIOGENETICS SWEDEN AB.

XX PI Pereira T, Poellinger L, Hellstroem M;

XX DR WPI; 2003-712876/67.

XX PT New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.

XX PS Example 8; Fig 22; 96pp; English.

XX CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. Sequences AAO23510-517 represent fragments within
 CC the N-TAD region of a murine HIF-1 alpha protein

XX SQ Sequence 15 AA;

Query Match 82.9%; Score 34; DB 7; Length 15;
 Best Local Similarity 87.5%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
 |||||


```

Db          2 MLAPYIPM 9

RESULT 12
AAO23485
ID AAO23485 standard; peptide; 15 AA.
XX AC AAO23485;
XX DT 12-FEB-2004 (first entry)
XX DE Murine HIF-1alpha protein N-TAD region fragment (residues 559-573).
XX KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
XX KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
XX KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
XX KW N-TAD.
XX OS Mus sp.
XX PN WO2003074560-A2.
XX PD 12-SEP-2003.
XX PF 05-MAR-2003; 2003WO-SE000372.
XX PR 05-MAR-2002; 2002US-0361333P.
XX PR 27-NOV-2002; 2002US-0429307P.
XX PA (ANGI-) ANGIOENTICS SWEDEN AB.
XX PI Pereira T, Poellinger L, Hellstroem M;
XX WPI; 2003-712876/67.
XX DR
XX PT New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
XX PT angiogenesis, or treating a condition associated with HIF-1alpha
XX PT underexpression in a cell, a group of cells, or an organism, e.g.
XX PT ischemia or inflammation.
XX PS Claim 39; Fig 22; 96pp; English.
XX CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
XX CC that has (a) an altered transactivation capacity and improved stability
XX CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
XX CC pharmaceutical composition are useful for increasing angiogenesis,
XX CC interfering with a normal response to reoxygenation following hypoxia, or
XX CC treating a condition associated with HIF-1alpha underexpression in a
XX CC cell, a group of cells, or an organism, e.g. ischemia, diabetic
XX CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
XX CC stroke. The proteins and pharmaceutical compositions are also useful for
XX CC mimicking the hypoxic response or artificially inducing a hypoxic
XX CC response in a cell, group of cells, or organism, inducing vascular
XX CC formation or vascular development in a cell or a group of cells,
XX CC increasing angiogenic activity in a cell, or influencing erythropoietin
XX CC production, metabolism, or an inflammatory response in a cell, a group of
XX CC cells, or an organism. Sequences AAO23482-489 represent fragments within
XX CC the N-TAD region of a murine HIF-1 alpha protein, that can be used for
XX CC treatment for hypoxic-related conditions
XX SQ Sequence 15 AA;

Query Match      82.9%; Score 34; DB 7; Length 15;
Best Local Similarity 87.5%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
   |||||
Db 2 MLAPYIPM 9

RESULT 13
AAE30163
ID AAE30163 standard; peptide; 16 AA.
XX AC AAE30163;
XX DT 24-FEB-2003 (first entry)
XX DE Peptide #3 used to block HIF-1alpha/pVHL interaction.
XX KW Hypoxia inducible factor; HIF hydroxylase; inflammatory disorder; cancer;
XX KW wound healing; ischaemia; transplantation; blood pressure; gene therapy.
XX OS Unidentified.
XX PN WO200274981-A2.
XX PD 26-SEP-2002.
XX PF 21-MAR-2002; 2002WO-GB001391.
XX PR 21-MAR-2001; 2001GB-00007123.
XX PR 02-AUG-2001; 2001GB-00018952.
XX PA (ISIS-) ISIS INNOVATION LTD.
XX PI Maxwell PH, Pugh CW, Ratcliffe PJ, Schofield CJ;
XX WPI; 2003-018808/01.
XX DR
XX PT Novel isolated polypeptide useful for treating ischemia, wound healing,
XX PT auto-, allo-, and xeno-transplantation, systemic high blood pressure,
XX PT cancer, or inflammatory disorders.
XX PS Example 1; Page 246; 256pp; English.
XX CC The invention relates to polypeptides having hypoxia inducible factor
XX CC (HIF) hydroxylase activity, referred to as PHD polypeptides (PHD 1,2 and
XX CC 3) and nucleic acid molecules encoding such polypeptides. Polypeptides of
XX CC the invention are used for treating conditions such as ischaemia, wound
XX CC healing, auto-, allo-, and xeno-transplantation, systemic high blood
XX CC pressure, cancer, or inflammatory disorders. They are useful in anti-
XX CC sense regulation of the HIF hydroxylase activity and in particular HIF
XX CC prolyl hydroxylase activity within a cell. They are also used to identify
XX CC additional substrates of HIF hydroxylases. Sequences of the invention are
XX CC used to design double stranded RNAs for use in RNA interference. They are
XX CC used as therapeutic agents and in purification, isolation, or screening
XX CC methods involving immuno-precipitation techniques and for detecting
XX CC polypeptides in biological samples. The invention is useful in gene
XX CC therapy. The present sequence is a peptide used to block HIF-1alpha/pVHL
XX CC interaction. This sequence is used in the invention
XX SQ Sequence 16 AA;

Query Match      82.9%; Score 34; DB 6; Length 16;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
   |||||
Db 3 MLAPYIPM 10

RESULT 14
AAO23488
ID AAO23488 standard; peptide; 16 AA.
XX AC AAO23488;
XX DT 12-FEB-2004 (first entry)
XX DE Murine HIF-1alpha protein N-TAD region fragment (residues 556-571).
XX KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
XX KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;

```

Wed Feb 9 06:57:54 2005

antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
N-TAD.

Mus sp.
WO2003074560-A2.
12-SEP-2003.
05-MAR-2003; 2003WO-SE000372.
05-MAR-2002; 2002US-0361333P.
27-NOV-2002; 2002US-0429307P.
(ANGI-) ANGIOGENETICS SWEDEN AB.
Pereira T, Poellinger L, Hellstrom M;
WPI; 2003-712876/67.
New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing angiogenesis, or treating a condition associated with HIF-1alpha underexpression in a cell, a group of cells, or an organism, e.g. ischemia or inflammation.

Claim 39; Fig 22; 96pp; English.
The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein that has (a) an altered transactivation capacity and improved stability at normoxia. The HIF-1alpha protein, polynucleotide, vector, and pharmaceutical composition are useful for increasing angiogenesis, interfering with a normal response to reoxygenation following hypoxia, or treating a condition associated with HIF-1alpha underexpression in a cell, a group of cells, or an organism, e.g. ischemia, diabetic retinopathy, inflammation, coronary heart disease, rheumatoid arthritis, stroke. The proteins and pharmaceutical compositions are also useful for mimicking the hypoxic response or artificially inducing a hypoxic response in a cell, group of cells, or organism, inducing vascular formation or vascular development in a cell, or influencing erythropoietin production, metabolism, or an inflammatory response in a cell, a group of cells, or an organism. Sequences AAO23482-489 represent fragments within the N-TAD region of a murine HIF-1 alpha protein, that can be used for treatment for hypoxic-related conditions

Sequence 16 AA;
Query Match 82.9%; Score 34; DB 7; Length 16;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIEM 8
Db 5 MLAPYIEM 12

RESULT 15
AAO23512
ID AAO23512 standard; peptide; 16 AA.
XX AAO23512;
XX AAO23512;
DT 12-FEB-2004 (first entry)
XX Murine HIF-1alpha protein N-TAD region fragment (residues 558-573).
XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW N-TAD.

Mus sp.

PN WO2003074560-A2.
XX 12-SEP-2003.
XX 05-MAR-2003; 2003WO-SE000372.
XX 05-MAR-2002; 2002US-0361333P.
PR 27-NOV-2002; 2002US-0429307P.
XX (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
XX
PI Pereira T, Poellinger L, Hellstrom M;
XX WPI; 2003-712876/67.
XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing angiogenesis, or treating a condition associated with HIF-1alpha underexpression in a cell, a group of cells, or an organism, e.g. ischemia or inflammation.

Example 8; Fig 22; 96pp; English.
The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein that has (a) an altered transactivation capacity and improved stability at normoxia. The HIF-1alpha protein, polynucleotide, vector, and pharmaceutical composition are useful for increasing angiogenesis, interfering with a normal response to reoxygenation following hypoxia, or treating a condition associated with HIF-1alpha underexpression in a cell, a group of cells, or an organism, e.g. ischemia, diabetic retinopathy, inflammation, coronary heart disease, rheumatoid arthritis, stroke. The proteins and pharmaceutical compositions are also useful for mimicking the hypoxic response or artificially inducing a hypoxic response in a cell, group of cells, or organism, inducing vascular formation or vascular development in a cell, or influencing erythropoietin production, metabolism, or an inflammatory response in a cell, a group of cells, or an organism. Sequences AAO23510-517 represent fragments within the N-TAD region of a murine HIF-1 alpha protein

Sequence 16 AA;

Query Match 82.9%; Score 34; DB 7; Length 16;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIEM 8
Db 3 MLAPYIEM 10

RESULT 16
AAO23484
ID AAO23484 standard; peptide; 16 AA.
XX AAO23484;
XX AAO23484;

DT 12-FEB-2004 (first entry)

XX Murine HIF-1alpha protein N-TAD region fragment (residues 558-573).
XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW N-TAD.

Mus sp.

XX WO2003074560-A2.
XX 12-SEP-2003.
XX 05-MAR-2003; 2003WO-SE000372.

PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.
PS Claim 39; Fig 22; 96pp; English.
XX
XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1alpha underexpression in a
CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell, or influencing erythropoietin
CC increasing angiogenic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. Sequences AAO233482-489 represent fragments within
CC the N-TAD region of a murine HIF-1 alpha protein, that can be used for
XX treatment for hypoxic-related conditions
XX
SQ Sequence 17 AA;
Query Match 82.9%; Score 34; DB 7; Length 17;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 MLAPTIPM 8
DB 4 MLAPVIPM 11
|||||
RESULT 19
AAO233511
ID AAO233511 standard; peptide; 17 AA.
XX
AC AAO233511;
XX
XX 12-FEB-2004 (first entry)
XX
XX Murine HIF-1alpha protein N-TAD region fragment (residues 5577573).
XX
XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW N-TAD.
XX
XX Mus sp.
XX
XX WO2003074560-A2.
XX
XX 12-SEP-2003.
XX
XX 05-MAR-2003; 2003WO-SE000372.
XX
XX 05-MAR-2002; 2002US-0361333P.
PR 27-NOV-2002; 2002US-0429307P.
XX
XX (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
XX Pereira T, Poellinger L, Hellstroem M;
XX
XX WPI; 2003-712876/67.
XX
XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.
XX
XX Example 8; Fig 22; 96pp; English.
PS
XX

CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1alpha underexpression in a
CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell, or influencing erythropoietin
CC increasing angiogenic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. Sequences AAO23510-517 represent fragments within
CC the N-TAD region of a murine HIF-1 alpha protein
XX
SQ Sequence 17 AA;
Query Match 82.9%; Score 34; DB 7; Length 17;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 MLAPTIPM 8
DB 4 MLAPVIPM 11
|||||
RESULT 20
ABP57669
ID ABP57669 standard; peptide; 18 AA.
XX
AC ABP57669;
XX
XX 30-APR-2003 (first entry)
XX
XX Hypoxia-inducible factor 1 alpha 557-574 peptide SEQ ID NO:1.
XX
XX Hypoxia-inducible factor 1 alpha; HIF-1 alpha; stabilisation; cytotoxic;
KW hypoxic; solid tumour; gene therapy; microbial fermentation; medicine;
KW tumour.
XX
XX Homo sapiens.
OS
XX WO200299104-A1.
XX
XX 12-DEC-2002.
XX
XX 04-JUN-2002; 2002WO-JP005482.
XX
XX 05-JUN-2001; 2001JP-00169948.
PR 05-JUN-2001; 2001JP-00169949.
XX
XX (POKK) POLA CHEM IND INC.
XX (HIRA/) HIRAOKA M.
XX (KOND/) KONDOH S.
XX
XX Hiraoka M, Kondoh S, Harada H;
XX
XX WPI; 2003-148670/14.
DR N-PSDB; ABZ71164.
XX
XX New DNA encoding a polypeptide imparting relative stability under hypoxic
PT conditions to proteins within the cell, useful for treatment of cancer
PT and improvement of microbial fermentation.
XX
XX Claim 1; Page 58; 144pp; Japanese.
PS
XX The present invention describes DNA encoding a hypoxia-inducible factor 1
CC alpha (HIF-1 alpha) amino acids 557 to 574 peptide (LDLEMLAPYIPMDDDFQL
CC see ABP57669) (I), or encoding a fusion protein containing at least 16
CC residues of (I), a nuclear localisation signal (NLS), and another
CC protein, and imparting relative stability under specific conditions of
XX

CC oxygen concentration within the cell. Also described: (1) vectors
 CC containing the DNA; (2) cells transformed by the vectors; (3) producing
 CC the fusion protein by culture of the transformed cells; (4) detecting
 CC hypoxic conditions in cells by monitoring the stability of the protein
 CC fused to (1) in cells transformed by vectors containing the DNA; (5)
 CC regulating the stability of proteins within the cell by transformation
 CC with the DNA; (6) inhibiting the development of cells under hypoxic
 CC conditions, using the fusion protein; (7) fusion proteins encoded by the
 CC DNA; and (8) stabilisation of cells under specific conditions of oxygen
 CC tension. (1) has cytostatic activity, and can be used for the
 CC stabilisation of a cytotoxic protein within cells in hypoxic regions of a
 CC solid tumour, and in gene therapy. (1) can be used in industrial
 CC microbial fermentation, and in medicine, especially in the treatment of
 CC tumours containing hypoxic regions
 XX
 SQ Sequence 18 AA;

Query Match 82.9%; Score 34; DB 6; Length 18;

Best Local Similarity 87.5%; Pred. No. 29;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8

Db 5 MLAPYIPM 12

RESULT 21

AAO23482

ID AAO23482 standard; peptide; 18 AA.

XX AC AAO23482;

DT 12-FEB-2004 (first entry)

DE Murine HIF-1alpha protein N-TAD region fragment (residues 556-573).

XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW N-TAD.

OS Mus sp.

PN WO2003074560-A2.

PD 12-SEP-2003.

XX PF 05-MAR-2003; 2003WO-SE000372.

XX PR 05-MAR-2002; 2002US-0361333P.

XX PR 27-NOV-2002; 2002US-0429307P.

XX PA (ANGI-) ANGIOGENETICS SWEDEN AB.

XX PI Pereira T, Poellinger L, Hellstroem M;

XX WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.

XX Claim 39; Fig 22; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,

CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. Sequences AAO23482-489 represent fragments within
 CC the N-TAD region of a murine HIF-1 alpha protein, that can be used for
 CC treatment for hypoxic-related conditions

XX SQ Sequence 18 AA;

Query Match 82.9%; Score 34; DB 7; Length 18;

Best Local Similarity 87.5%; Pred. No. 29;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8

Db 5 MLAPYIPM 12

RESULT 22

AAO23510

ID AAO23510 standard; peptide; 18 AA.

XX AC AAO23510;

DT 12-FEB-2004 (first entry)

DE Murine HIF-1alpha protein N-TAD region fragment (residues 556-573).
 XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW N-TAD.

OS Mus sp.

PN WO2003074560-A2.

PD 12-SEP-2003.

XX PF 05-MAR-2003; 2003WO-SE000372.

XX PR 05-MAR-2002; 2002US-0361333P.

XX PR 27-NOV-2002; 2002US-0429307P.

XX PA (ANGI-) ANGIOGENETICS SWEDEN AB.

XX PI Pereira T, Poellinger L, Hellstroem M;

XX WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.

XX Example 8; Fig 22; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin

CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. Sequences AAO23510-517 represent fragments within
 CC the N-TAD region of a murine HIF-1 alpha protein
 XX
 SQ Sequence 18 AA;

Query Match 82.9%; Score 34; DB 7; Length 18;
 Best Local Similarity 87.5%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 |||||
 Db 5 MLAPYIPM 12

RESULT 23
 AAB49912
 ID AAB49912 standard; peptide; 19 AA.
 XX
 AC AAB49912;
 XX
 DT 06-MAR-2001 (first entry)
 XX
 DE Human/murine HIF-1alpha subunit conserved motif #8.
 XX
 KW Mouse; human; HIF-1alpha; von Hippel-Lindau syndrome protein; VHL;
 KW hypoxia inducible factor-1; cancer; ischaemia.
 XX
 OS Mus sp.
 OS Homo sapiens.
 XX
 PN WO200069908-A1.
 XX
 PD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-GB001826.
 XX
 PR 12-MAY-1999; 99GB-00011047.
 XX
 PA (ISIS-) ISIS INNOVATION LTD.
 PI Ratcliffe PJ, Maxwell PH, Pugh CW;
 XX WPI; 2001-025006/03.
 DR
 XX Assaying for von Hippel Lindau (VHL)-hypoxia inducible factor (HIF) alpha
 PT subunit interaction modulators for treating ischemia by contacting a VHL
 PT protein and an HIF subunit protein with a putative modulator.
 XX
 PS Claim 13; Page 49; 56pp; English.

CC The present invention describes a novel assay for use in identifying
 CC modulators of the von Hippel-Lindau protein (VHL) and hypoxia inducible
 CC factor-1 alpha subunit (HIF-1alpha) interaction. The assay comprises
 CC contacting the VHL protein, the HIF-1alpha subunit and the putative
 CC modulator under conditions where the former two would normally complex.
 CC Modulators of this type are useful in the treatment of cancer and
 CC ischaemic conditions such as coronary, cerebral and vascular
 CC insufficiency
 XX
 SQ Sequence 19 AA;

Query Match 82.9%; Score 34; DB 4; Length 19;
 Best Local Similarity 87.5%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 |||||
 Db 6 MLAPYIPM 13

RESULT 24
 AAE30167

ID AAE30167 standard; peptide; 19 AA.
 XX
 AC AAE30167;
 XX
 DT 24-FEB-2003 (first entry)
 XX
 DE Peptide #7 used to block HIF-1alpha/pVHL interaction.
 XX
 KW Hypoxia inducible factor; HIF hydroxylase; inflammatory disorder; cancer;
 KW wound healing; ischaemia; transplantation; blood pressure; gene therapy.
 XX
 OS Unidentified.
 XX
 PN WO200274981-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 21-MAR-2002; 2002WO-GB001381.
 XX
 PR 21-MAR-2001; 2001GB-00007123.
 PR 02-AUG-2001; 2001GB-00018952.
 XX
 PA (ISIS-) ISIS INNOVATION LTD.
 XX
 PI Maxwell PH, Pugh CW, Ratcliffe PJ, Schofield CJ;
 XX WPI; 2003-018808/01.
 DR
 XX Novel isolated polypeptide useful for treating ischemia, wound healing,
 PT auto-, allo-, and xeno-transplantation, systemic high blood pressure,
 PT cancer, or inflammatory disorders.
 XX
 PS Example 1; Page 247; 256pp; English.

CC The invention relates to polypeptides having hypoxia inducible factor
 CC (HIF) hydroxylase activity, referred to as PHD polypeptides (PHD 1,2 and
 CC 3) and nucleic acid molecules encoding such polypeptides. Polypeptides of
 CC the invention are used for treating conditions such as ischaemia, wound
 CC healing, auto-, allo-, and xeno-transplantation, systemic high blood
 CC pressure, cancer, or inflammatory disorders. They are useful in anti-
 CC sense regulation of the HIF hydroxylase activity and in particular HIF
 CC prolyl hydroxylase activity within a cell. They are also used to identify
 CC additional substrates of HIF hydroxylases. Sequences of the invention are
 CC used to design double stranded RNAs for use in RNA interference. They are
 CC used as therapeutic agents and in purification, isolation, or screening
 CC methods involving immuno-precipitation techniques and for detecting
 CC polypeptides in biological samples. The invention is useful in gene
 CC therapy. The present sequence is a peptide used to block HIF-1alpha/pVHL
 CC interaction. This sequence is used in the invention

Query Match 82.9%; Score 34; DB 6; Length 19;
 Best Local Similarity 87.5%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 |||||
 Db 6 MLAPYIPM 13

RESULT 25
 AAE30144
 ID AAE30144 standard; peptide; 19 AA.
 XX
 AC AAE30144;
 XX
 DT 24-FEB-2003 (first entry)
 XX
 DE HIF1alpha subunit antagonist #1.
 XX
 KW Hypoxia inducible factor; HIF hydroxylase; inflammatory disorder; cancer;
 KW wound healing; ischaemia; transplantation; blood pressure; gene therapy;

KW antagonist.
 XX OS Unidentified.
 XX FH Key Location/Qualifiers
 FT Modified-site 2
 FT /label= Hyp
 XX WO200274981-A2.
 XX PN 26-SEP-2002.
 XX PD 21-MAR-2002; 2002WO-GB001381.
 XX PF 21-MAR-2001; 2001GB-00007123.
 XX PR 02-AUG-2001; 2001GB-00018952.
 XX PA (ISIS-) ISIS INNOVATION LTD.
 XX PI Maxwell PH, Pugh CW, Ratcliffe PJ, Schofield CJ;
 XX WPI; 2003-018808/01.
 XX Novel isolated polypeptide useful for treating ischemia, wound healing,
 PT auto-, allo-, and xeno-transplantation, systemic high blood pressure,
 PT cancer, or inflammatory disorders.
 XX Claim 49; Page 196; 256pp; English.
 XX The invention relates to polypeptides having hypoxia inducible factor
 CC (HIF) hydroxylase activity, referred to as PHD polypeptides (PHD 1,2 and
 CC 3) and nucleic acid molecules encoding such polypeptides. Polypeptides of
 CC the invention are used for treating conditions such as ischaemia, wound
 CC healing, auto-, allo-, and xeno-transplantation, systemic high blood
 CC pressure, cancer, or inflammatory disorders. They are useful in anti-
 CC sense regulation of the HIF hydroxylase activity and in particular HIF
 CC prolyl hydroxylase activity within a cell. They are also used to identify
 CC additional substrates of HIF hydroxylases. Sequences of the invention are
 CC used as therapeutic agents and in purification, isolation, or screening
 CC methods involving immuno-precipitation techniques and for detecting
 CC polypeptides in biological samples. The invention is useful in gene
 CC therapy. The present sequence is HIFalpha subunit antagonist. This
 CC sequence is used in the invention
 XX Sequence 19 AA;
 SQ
 Query Match 82.9%; Score 34; DB 6; Length 19;
 Best Local Similarity 87.5%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPYIPM 8
 DB ||||| ||||
 6 MLAPYIPM 13
 RESULT 26
 AAE30162
 ID AAE30162 standard; peptide; 19 AA.
 XX AC AAE30162;
 XX DT 24-FEB-2003 (first entry)
 XX DE Peptide #2 used to block HIF-1alpha/pVHL interaction.
 XX KW Hypoxia inducible factor; HIF hydroxylase; inflammatory disorder; cancer;
 KW wound healing; ischaemia; transplantation; blood pressure; gene therapy.
 XX OS Unidentified.
 XX PN WO200274981-A2.

PD 26-SEP-2002.
 XX 21-MAR-2002; 2002WO-GB001381.
 XX PR 21-MAR-2001; 2001GB-00007123.
 PR 02-AUG-2001; 2001GB-00018952.
 XX (ISIS-) ISIS INNOVATION LTD.
 XX PI Maxwell PH, Pugh CW, Ratcliffe PJ, Schofield CJ;
 XX WPI; 2003-018808/01.
 XX Novel isolated polypeptide useful for treating ischemia, wound healing,
 PT auto-, allo-, and xeno-transplantation, systemic high blood pressure,
 PT cancer, or inflammatory disorders.
 XX Example 1; Page 246; 256pp; English.
 XX The invention relates to polypeptides having hypoxia inducible factor
 CC (HIF) hydroxylase activity, referred to as PHD polypeptides (PHD 1,2 and
 CC 3) and nucleic acid molecules encoding such polypeptides. Polypeptides of
 CC the invention are used for treating conditions such as ischaemia, wound
 CC healing, auto-, allo-, and xeno-transplantation, systemic high blood
 CC pressure, cancer, or inflammatory disorders. They are useful in anti-
 CC sense regulation of the HIF hydroxylase activity and in particular HIF
 CC prolyl hydroxylase activity within a cell. They are also used to identify
 CC additional substrates of HIF hydroxylases. Sequences of the invention are
 CC used as therapeutic agents and in purification, isolation, or screening
 CC methods involving immuno-precipitation techniques and for detecting
 CC polypeptides in biological samples. The invention is useful in gene
 CC therapy. The present sequence is a peptide used to block HIF-1alpha/pVHL
 CC interaction. This sequence is used in the invention
 XX Sequence 19 AA;
 SQ
 Query Match 82.9%; Score 34; DB 6; Length 19;
 Best Local Similarity 87.5%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPYIPM 8
 DB ||||| ||||
 6 MLAPYIPM 13
 RESULT 27
 AAE30172
 ID AAE30172 standard; peptide; 19 AA.
 XX AC AAE30172;
 XX DT 24-FEB-2003 (first entry)
 XX DE Human HIF1-alpha peptide #2.
 XX KW Hypoxia inducible factor; HIF hydroxylase; inflammatory disorder; cancer;
 KW wound healing; ischaemia; transplantation; blood pressure; gene therapy;
 KW human; HIF1-alpha.
 XX OS Homo sapiens.
 XX PN WO200274981-A2.
 XX PD 26-SEP-2002.
 XX 21-MAR-2002; 2002WO-GB001381.
 XX PR 21-MAR-2001; 2001GB-00007123.
 PR 02-AUG-2001; 2001GB-00018952.
 XX (ISIS-) ISIS INNOVATION LTD.
 XX PA

PI Maxwell PH, Pugh CW, Ratcliffe PJ, Schofield CJ;
 XX WPI; 2003-018808/01.
 XX Novel isolated polypeptide useful for treating ischemia, wound healing,
 PT auto-, allo-, and xeno-transplantation, systemic high blood pressure,
 PT cancer, or inflammatory disorders.
 XX
 XX Disclosure; Page 252; 256pp; English.
 XX
 XX The invention relates to polypeptides having hypoxia inducible factor
 CC (HIF) hydroxylase activity, referred to as PHD polypeptides (PHD 1,2 and
 CC 3) and nucleic acid molecules encoding such polypeptides. Polypeptides of
 CC the invention are used for treating conditions such as ischaemia, wound
 CC healing, auto-, allo-, and xeno-transplantation, systemic high blood
 CC pressure, cancer, or inflammatory disorders. They are useful in anti-
 CC sense regulation of the HIF hydroxylase activity and in particular HIF
 CC prolyl hydroxylase activity within a cell. They are also used to identify
 CC additional substrates of HIF hydroxylases. Sequences of the invention are
 CC used to design double stranded RNAs for use in RNA interference. They are
 CC used as therapeutic agents and in purification, isolation, or screening
 CC methods involving immuno-precipitation techniques and for detecting
 CC polypeptides in biological samples. The invention is useful in gene
 CC therapy. The present sequence is human HIF1-alpha peptide. This sequence
 CC is used in the invention
 XX
 XX Sequence 19 AA;
 SQ

Query Match 82.9%; Score 34; DB 6; Length 19;
 Best Local Similarity 87.5%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 DB 6 MLAPYIPM 13
 |||||
 |||||

RESULT 28
 AAE30158
 ID AAE30158 standard; peptide; 19 AA.
 XX
 AC AAE30158;
 DT 24-FEB-2003 (first entry)
 XX
 DE HIF-1alpha pVHL minimal binding domain.
 XX
 KW Hypoxia inducible factor; HIF hydroxylase; inflammatory disorder; cancer;
 KW wound healing; ischaemia; transplantation; blood pressure; gene therapy;
 KW human.
 XX
 OS Homo sapiens.
 XX
 FN WO200274981-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 21-MAR-2002; 2002WO-GB001381.
 XX
 PR 21-MAR-2001; 2001GB-00007123.
 PR 02-AUG-2001; 2001GB-00018952.
 XX
 PA (ISIS-) ISIS INNOVATION LTD.
 XX
 PI Maxwell PH, Pugh CW, Ratcliffe PJ, Schofield CJ;
 XX WPI; 2003-018808/01.
 XX Novel isolated polypeptide useful for treating ischemia, wound healing,
 PT auto-, allo-, and xeno-transplantation, systemic high blood pressure,
 PT cancer, or inflammatory disorders.
 XX
 XX Example 1; Page 245; 256pp; English.

XX The invention relates to polypeptides having hypoxia inducible factor
 CC (HIF) hydroxylase activity, referred to as PHD polypeptides (PHD 1,2 and
 CC 3) and nucleic acid molecules encoding such polypeptides. Polypeptides of
 CC the invention are used for treating conditions such as ischaemia, wound
 CC healing, auto-, allo-, and xeno-transplantation, systemic high blood
 CC pressure, cancer, or inflammatory disorders. They are useful in anti-
 CC sense regulation of the HIF hydroxylase activity and in particular HIF
 CC prolyl hydroxylase activity within a cell. They are also used to identify
 CC additional substrates of HIF hydroxylases. Sequences of the invention are
 CC used to design double stranded RNAs for use in RNA interference. They are
 CC used as therapeutic agents and in purification, isolation, or screening
 CC methods involving immuno-precipitation techniques and for detecting
 CC polypeptides in biological samples. The invention is useful in gene
 CC therapy. The present sequence is human HIF1-alpha peptide. This sequence
 CC is used in the invention
 XX
 XX Sequence 19 AA;
 SQ

Query Match 82.9%; Score 34; DB 6; Length 19;
 Best Local Similarity 87.5%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 DB 6 MLAPYIPM 13
 |||||
 |||||

RESULT 29
 ABR82378
 ID ABR82378 standard; peptide; 19 AA.
 XX
 AC ABR82378;
 DT 06-NOV-2003 (first entry)
 XX
 DE Hypoxia-inducible factor 1 (HIF-1) alpha peptide inhibitor.
 XX
 KW HIF-1; hypoxia-inducible factor 1; HIF-1 alpha; ubiquitination; EPO;
 KW erythropoietin; vascular endothelial growth factor; VEGF; glycolytic;
 KW tranquilizer; vulnerary; cardiant; cerebroprotective; angiogenesis.
 XX
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Modified-site 9
 FT /label= Hyp
 FT /note= "hydroxyproline"
 XX
 FN WO2003057820-A2.
 XX
 PD 17-JUL-2003.
 XX
 PF 04-OCT-2002; 2002WO-US031699.
 XX
 PR 21-DEC-2001; 2001US-00032361.
 XX
 PA (KIMB) KIMBERLY-CLARK WORLDWIDE INC.
 XX
 FI McGrath K;
 XX
 DR WPI; 2003-645988/61.
 XX
 PT Novel peptide inhibitor of hypoxia-inducible factor 1 alpha
 PT ubiquitination, and activator of vascular endothelial growth factor
 PT transcription useful for treating tissue injuries including wounds,
 PT surgical incisions.
 XX
 PS Claim 2; Page 8; 37pp; English.
 XX
 CC The invention relates to peptide inhibitors of hypoxia-inducible factor
 CC (HIF-1) alpha ubiquitination. The peptide inhibitors thereby activate the
 CC transcription of erythropoietin (EPO), vascular endothelial growth factor

CC (VEGF) and certain glycolytic enzymes. The peptide inhibitors are useful
 CC for treating tissue injuries including wounds, surgical incisions,
 CC chronic wounds, heart disease and stroke. The present sequence represents
 CC a specific example of HIF-1 alpha peptide inhibitor, containing the
 CC oxygen-dependent degradation sequence of HIF-1 alpha
 XX
 SQ Sequence 19 AA;

Query Match 82.9%; Score 34; DB 6; Length 19;
 Best Local Similarity 87.5%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
 |||||
 Db 6 MLAPYIPM 13

RESULT 30

ADP56728
 ID ADP56728 standard; peptide; 19 AA.

XX AC ADP56728;

DT 09-SEP-2004 (first entry)

XX DE Substrate peptide used in human HIF prolyl hydroxylase screening assay.
 XX KW fat metabolism; HIF1alpha; hypoxia inducible factor alpha subunit;
 KW atherosclerosis; diabetes; obesity; HIF prolyl hydroxylase substrate;
 KW human; HIF-PH.

XX OS Homo sapiens.

XX PN WO2004052285-A2.

XX PD 24-JUN-2004.

XX PF 05-DEC-2003; 2003WO-US038690.

XX PR 06-DEC-2002; 2002US-0431351P.

XX PR 06-JUN-2003; 2003US-0476331P.

XX PR 06-JUN-2003; 2003US-0476726P.

XX PR 04-DEC-2003; 2003US-00729167.

XX PA (FIBR-) FIBROGEN INC.

XX PI Fournay PD, Guenzler-Pukall V, Klaus SJ, Lin AY, Neff TB;

PI Seeley TW;

XX WPI; 2004-468689/44.

XX DR Regulating fat metabolism or fat metabolic process in subjects, by
 PT stabilizing human foreskin fibroblasts alpha in subject, thus regulating
 PT fat metabolism or fat metabolic process in subject.

XX PS Example 9; SEQ ID NO 1; 66pp; English.

XX CC The invention relates to a novel method for regulating fat metabolism or
 CC the fat metabolic process in a subject which comprises stabilising human
 CC foreskin fibroblast HIF1alpha (hypoxia inducible factor alpha subunit) in
 CC the subject, or administering a compound that inhibits HIF hydroxylase
 CC activity, thus regulating fat metabolism or the fat metabolic process in
 CC the subject. The method of the invention may be useful for regulating fat
 CC metabolism or a fat metabolic process in a subject. The subject is an
 CC animal, preferably a mammal, more preferably human and the method is
 CC performed in a human cell, tissue or organ. The method may be useful for
 CC treating or preventing atherosclerosis, diabetes and obesity in a
 CC subject. The current sequence is that of the substrate peptide of the
 CC invention which is used during a screening assay of human HIF prolyl
 CC hydroxylase (HIF-PH).

XX SQ Sequence 19 AA;

Query Match

Best Local Similarity 82.9%; Score 34; DB 8; Length 19;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8
 |||||
 Db 6 MLAPYIPM 13

RESULT 31

ADP79479
 ID ADP79479 standard; peptide; 19 AA.

XX AC ADP79479;

XX DT 04-NOV-2004 (first entry)

XX DE Hypoxia inducible factor prolyl hydroxylase substrate peptide.

XX KW Human; Hypoxia inducible factor prolyl hydroxylase; glucose metabolism;
 KW antidiabetic; anorectic; hypotensive; antilipemic; nephrotropic;
 KW neuroprotective; ophthalmological; antiarteriosclerotic; vasotropic;
 KW enzyme.

XX OS Homo sapiens.

XX PN WO2004052284-A2.

XX PD 24-JUN-2004.

XX PF 05-DEC-2003; 2003WO-US038689.

XX PR 06-DEC-2002; 2002US-0431351P.

XX PR 06-JUN-2003; 2003US-0476331P.

XX PR 06-JUN-2003; 2003US-0476726P.

XX PR 04-DEC-2003; 2003US-00729704.

XX PA (FIBR-) FIBROGEN INC.

XX PI Guenzler-Pukall V, Klaus SJ, Langsetmo Parobok I, Seeley TW;

XX WPI; 2004-468688/44.

XX DR Regulating glucose metabolism or glucose metabolic process in subject,
 PT involves stabilizing hypoxia inducible factor alpha in subject, or
 PT administering to subject compound inhibiting hypoxia inducible factor
 PT hydroxylase activity.

XX PS Example 14; SEQ ID NO 5; 74pp; English.

XX CC The present sequence is that of a substrate peptide for hypoxia inducible
 CC factor (HIF) prolyl hydroxylase. It was used in an example from the
 CC invention for the identification of compounds useful for HIF alpha
 CC stabilisation. The invention provides methods and compounds for
 CC regulating glucose metabolism by stabilising HIF alpha, especially by
 CC administering a compound that inhibits HIF hydroxylase activity. The
 CC method of stabilising HIF alpha is used in claimed methods for achieving
 CC glucose homeostasis, decreasing blood glucose levels, decreasing glycated
 CC haemoglobin levels, altering expression of a glucose regulatory factor,
 CC altering expression of a glycolytic factor, treating or preventing
 CC diabetes, treating or preventing a disorder associated with increased
 CC blood glucose levels (especially diabetes, hyperglycaemia, obesity,
 CC hypertension, hyperlipidaemia, nephropathy, neuropathy, retinopathy,
 CC impaired glucose tolerance, atherosclerosis and vascular disease),
 CC treating or preventing a condition associated with diabetes, decreasing
 CC blood triglyceride levels, reducing insulin resistance, and increasing
 CC glycaemic control in a subject.

XX SQ Sequence 19 AA;

Query Match

Best Local Similarity 82.9%; Score 34; DB 8; Length 19;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
 DB 6 MLAPYIPM 13

RESULT 32
 ABP55440
 ID ABP55440 standard; peptide; 20 AA.
 XX
 AC ABP55440;
 XX
 DT 05-FEB-2003 (first entry)
 XX
 DE Hypoxia-inducible factor (HIF) 1 alpha peptide.
 XX
 KW Hypoxia; hypoxia-inducible factor; HIF1-alpha; hypoxic-related disorder;
 KW ischaemic-related disorder; hypoxia-inducible factor-related disorder;
 KW prolyl hydroxylation; HIF; hypoxic; ischaemic; vasotrophic; cardiatic;
 KW cerebroprotective; cytotatic; thrombolytic; antidiabetic; nephrotropic;
 KW myocardial infarction; heart disease; stroke; cancer; diabetes;
 KW cell-proliferating disorder; deep vein thrombosis; pulmonary embolus;
 KW renal failure; angiogenesis; vascularisation; prolyl hydroxylase.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 9
 FT /label= hydroxyproline
 XX
 PN WO200274249-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 20-MAR-2002; 2003WO-US008886.
 XX
 PR 20-MAR-2001; 2001US-0277425P.
 PR 20-MAR-2001; 2001US-0277431P.
 PR 20-MAR-2001; 2001US-0277440P.
 PR 09-NOV-2001; 2001US-0323234P.
 PR 09-NOV-2001; 2001US-0332493P.
 PR 09-NOV-2001; 2001US-0345200P.
 PR 20-DEC-2001; 2001US-0345259P.
 PR 20-DEC-2001; 2001US-0345131P.
 PR 20-DEC-2001; 2001US-0345132P.
 PR 19-MAR-2002; 2002US-00101812.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Kaelin WG, Ivan M;
 XX
 DR WPI; 2003-058330/05.
 XX
 PT Treating or preventing a hypoxic- or ischemic-related disorder, e.g.
 PT myocardial infarction, stroke, cancer, thrombosis or renal failure, by
 PT administering a modulator prolyl hydroxylation of hypoxia-inducible
 PT factor (HIF).
 XX
 PS Disclosure; Page 26; 128pp; English.
 XX
 CC The present invention describes a method for treating or preventing a
 CC hypoxic-related disorder, ischaemic-related disorder, or hypoxia-
 CC inducible factor (HIF)-related disorder in a subject by administering to
 CC the subject a compound that modulates prolyl hydroxylation of HIF, such
 CC that the hypoxic-, ischaemic-, or HIF-related disorder is treated,
 CC prevented, reversed or stabilised. HIF related sequences can have
 CC vasotrophic, cardiatic, cerebroprotective, cytotatic, thrombolytic,
 CC antidiabetic, and nephrotropic activities. The method is useful for
 CC treating hypoxia-related disorder, ischaemic-related disorder or HIF-
 CC related disorder. In particular, the hypoxic- or ischaemic-related
 CC disorder is an acute event (e.g. myocardial infarction, heart disease,
 CC stroke, cancer or cell-proliferating disorder, or diabetes) or a chronic
 CC event (e.g. deep vein thrombosis, pulmonary embolus or renal failure),

CC especially a chronic event not caused by tissue scarring. The method is
 CC also useful for increasing angiogenesis or vascularisation. The present
 CC sequence represents a human hypoxia-inducible factor 1 alpha (HIF1-alpha)
 CC peptide which is given in the exemplification of the present invention
 XX
 SQ Sequence 20 AA;
 Query Match 82.9%; Score 34; DB 6; Length 20;
 Best Local Similarity 87.5%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 DB 6 MLAPYIPM 13

RESULT 33
 ADO22337
 ID ADO22337 standard; peptide; 20 AA.
 XX
 AC ADO22337;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE HIF-1alpha ligand binding site peptide SEQ ID NO:15.
 XX
 KW transgenic; transgenic non-human animal; light-generating fusion protein;
 KW ligand binding site; light-generating polypeptide moiety;
 KW hypoxia-inducible factor 1 alpha; HIF-1alpha; hypoxic tissue;
 KW respiratory; cytotatic; vasotrophic; virucide; hypoxic condition; cancer;
 KW ischaemia; viral infection; drug screening; drug discovery;
 KW ligand binding site peptide.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO2004042361-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 03-NOV-2003; 2003WO-US035154.
 XX
 PR 04-NOV-2002; 2002US-00287670.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Kaelin WG, Livingston DM, Kim T;
 XX
 DR WPI; 2004-400725/37.
 XX
 PT New transgenic non-human animal comprising light-generating fusion
 PT protein, useful in diagnosing and treating hypoxic conditions, cancer,
 PT ischemia and viral infections and in drug screening and discovery.
 XX
 PS Disclosure; SEQ ID NO 15; 111pp; English.
 XX
 CC The present invention describes a transgenic non-human animal comprising
 CC a recombinant nucleic acid molecule stably integrated into the genome of
 CC the animal, where the molecule encodes a light-generating fusion protein
 CC comprising a ligand binding site and a light-generating polypeptide
 CC moiety. Also described: (1) an isolated cell of the animal; (2) producing
 CC a transgenic non-human animal; (3) identifying a compound capable of
 CC modifying an activity of hypoxia-inducible factor (HIF) 1 alpha; and (4)
 CC detecting hypoxic tissue. The compound has respiratory, cytotatic,
 CC vasotrophic and virucide activities. The transgenic non-human animal,
 CC light-generating fusion protein, agents, kits and compositions are useful
 CC in diagnosing and treating hypoxic conditions, cancer, ischaemia and
 CC viral infections and in drug screening and discovery. The present
 CC sequence represents a HIF-1alpha ligand binding site peptide, which is
 CC used in the exemplification of the present invention.
 XX
 SQ Sequence 20 AA;

```
Query Match      82.9%; Score 34; DB 8; Length 20;
Best Local Similarity 87.5%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAPTIPM 8
Db      6 MLAPYIPM 13
      |||||
      |||||

RESULT 34
AAO23532
ID AAO23532 standard; peptide; 28 AA.
XX
AC AAO23532;
XX
DT 12-FEB-2004 (first entry)
XX
DE Fluorescein HIV-TAT cellular uptake signal-HIF-1alpha peptide 2.
XX
KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; HIV; tat;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy.
XX
OS Synthetic.
XX
FH Key
FT Peptide      1..13
FT Modified-site /note= "HIV-TAT cellular uptake signal peptide"
FT FT           /note= "5'- fluorescein label"
FT FT           14..28
FT FT           /note= "HIF-1alpha mutant peptide"
XX
PN WO2003074560-A2.
XX
PD 12-SEP-2003.
XX
PF 05-MAR-2003; 2003WO-SE000372.
XX
PR 05-MAR-2002; 2002US-0361333P.
XX
PT 27-NOV-2002; 2002US-0429307P.
XX
PA (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
PI Pereira T, Poellinger L, Hellstroem M;
XX
DR WPI; 2003-712876/67.
XX
PS New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
XX PT angiogenesis, or treating a condition associated with HIF-1alpha
XX PT underexpression in a cell, a group of cells, or an organism, e.g.
XX PT ischemia or inflammation.
XX
PS Example 16; Fig 32; 96pp; English.
XX
CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
XX CC that has (a) an altered transactivation capacity and improved stability
XX CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
XX CC pharmaceutical composition are useful for increasing angiogenesis,
XX CC interfering with a normal response to reoxygenation following hypoxia, or
XX CC treating a condition associated with HIF-1alpha underexpression in a
XX CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
XX CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
XX CC stroke. The proteins and pharmaceutical compositions are also useful for
XX CC mimicking the hypoxic response or artificially inducing a hypoxic
XX CC response in a cell, group of cells, or organism, inducing vascular
XX CC formation or vascular development in a cell or a group of cells.
XX CC increasing angiogenic activity in a cell, or influencing erythropoietin
XX CC production, metabolism, or an inflammatory response in a cell, a group of
XX CC cells, or an organism. The present sequence represents a HIF-1alpha
XX CC mutant peptide linked to a fluorescein labeled HIV-TAT cellular uptake
XX CC signal peptide

SQ      Sequence 28 AA;
Query Match      82.9%; Score 34; DB 7; Length 29;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAPTIPM 8
Db      15 MLAPYIPM 22
      |||||
      |||||

RESULT 35
AAO23475
ID AAO23475 standard; peptide; 29 AA.
XX
AC AAO23475;
XX
DT 12-FEB-2004 (first entry)
XX
DE Murine HIF-1alpha protein mutant fragment.
XX
KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW mutant.
XX
OS Mus sp.
XX
PN WO2003074560-A2.
XX
PD 12-SEP-2003.
XX
PF 05-MAR-2003; 2003WO-SE000372.
XX
PR 05-MAR-2002; 2002US-0361333P.
XX
PT 27-NOV-2002; 2002US-0429307P.
XX
PA (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
PI Pereira T, Poellinger L, Hellstroem M;
XX
DR WPI; 2003-712876/67.
XX
PS New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
XX PT angiogenesis, or treating a condition associated with HIF-1alpha
XX PT underexpression in a cell, a group of cells, or an organism, e.g.
XX PT ischemia or inflammation.
XX
PS Claim 39; Page 90; 96pp; English.
XX
CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
XX CC that has (a) an altered transactivation capacity and improved stability
XX CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
XX CC pharmaceutical composition are useful for increasing angiogenesis,
XX CC interfering with a normal response to reoxygenation following hypoxia, or
XX CC treating a condition associated with HIF-1alpha underexpression in a
XX CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
XX CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
XX CC stroke. The proteins and pharmaceutical compositions are also useful for
XX CC mimicking the hypoxic response or artificially inducing a hypoxic
XX CC response in a cell, group of cells, or organism, inducing vascular
XX CC formation or vascular development in a cell or a group of cells.
XX CC increasing angiogenic activity in a cell, or influencing erythropoietin
XX CC production, metabolism, or an inflammatory response in a cell, a group of
XX CC cells, or an organism. The present sequence represents a specific example
XX CC of a murine HIF-1 alpha mutant fragment used for treatment for hypoxic-
XX CC related conditions
XX
SQ      Sequence 29 AA;
```

```

Qy      1 MLAPTIEM 8
Db      15 MLAPYIPM 22

RESULT 36
AAO23478
ID AAO23478 standard; peptide; 29 AA.
XX
AC AAO23478;
XX
DT 12-FEB-2004 (first entry)
XX
DE Murine HIF-1alpha protein mutant fragment.
XX
KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW mutant.
XX
OS Mus sp.
XX
PN WO2003074560-A2.
XX
PD 12-SEP-2003.
XX
PF 05-MAR-2003; 2003WO-SE000372.
XX
PR 05-MAR-2002; 2002US-0361333P.
XX
PR 27-NOV-2002; 2002US-0429307P.
XX
PA (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
PI Pereira T, Poellinger L, Hellstroem M;
XX
DR WPI; 2003-712876/67.
XX
PT New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.
XX
PS Claim 39; Page 92; 96pp; English.
XX
CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1alpha underexpression in a
CC cell, a group of cells, or an organism, e.g. ischemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell or a group of cells,
CC increasing angiogenic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. The present sequence represents a specific example
CC of a murine HIF-1 alpha mutant fragment used for treatment for hypoxic-
CC related conditions
XX
SQ Sequence 29 AA;

Query Match      82.9%; Score 34; DB 7; Length 29;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MLAPTIEM 8
Db      15 MLAPYIPM 22

RESULT 37
AAO23476
ID AAO23476 standard; peptide; 29 AA.
XX
AC AAO23476;
XX
DT 12-FEB-2004 (first entry)
XX
DE Murine HIF-1alpha protein mutant fragment.
XX
KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW mutant.
XX
OS Mus sp.
XX
PN WO2003074560-A2.
XX
PD 12-SEP-2003.
XX
PF 05-MAR-2003; 2003WO-SE000372.
XX
PR 05-MAR-2002; 2002US-0361333P.
XX
PR 27-NOV-2002; 2002US-0429307P.
XX
PA (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
PI Pereira T, Poellinger L, Hellstroem M;
XX
DR WPI; 2003-712876/67.
XX
PT New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.
XX
PS Claim 39; Page 91; 96pp; English.
XX
CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1alpha underexpression in a
CC cell, a group of cells, or an organism, e.g. ischemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell or a group of cells,
CC increasing angiogenic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. The present sequence represents a specific example
CC of a murine HIF-1 alpha mutant fragment used for treatment for hypoxic-
CC related conditions
XX
SQ Sequence 29 AA;

Query Match      82.9%; Score 34; DB 7; Length 29;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 MLAPTIEM 8
Db      15 MLAPYIPM 22

RESULT 38
AAO23503
ID AAO23503 standard; peptide; 29 AA.
XX

```

AC AAO23503;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Murine HIF-1alpha protein N-TAD region mutant fragment D568A.
 XX
 KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW N-TAD; mutant.
 XX
 OS Mus sp.
 XX
 PN WO2003074560-A2.
 XX
 PD 12-SEP-2003.
 XX
 PF 05-MAR-2003; 2003WO-SE000372.
 XX
 PR 05-MAR-2002; 2002US-0361333P.
 PR 27-NOV-2002; 2002US-0429307P.
 XX
 PA (ANGI-) ANGIOGENETICS SWEDEN AB.
 XX
 PI Pereira T, Poellinger L, Hellstroem M;
 XX WPI; 2003-712876/67.
 XX
 PT New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.
 XX
 PS Example 6; Fig 19; 96pp; English.
 XX
 CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. Sequences AAO23500-509represent mutant fragments
 CC within the N-TAD region of a murine HIF-1 alpha protein
 XX
 SQ Sequence 29 AA;
 Query Match 82.9%; Score 34; DB 7; Length 29;
 Best Local Similarity 87.5%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 DB 15 MLAPYIPM 22
 RESULT 39
 AAO23506
 ID AAO23506 standard; peptide; 29 AA.
 XX
 AC AAO23506;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Murine HIF-1alpha protein N-TAD region mutant fragment F571A.
 XX

KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW N-TAD; mutant.
 XX
 OS Mus sp.
 XX
 PN WO2003074560-A2.
 XX
 PD 12-SEP-2003.
 XX
 PF 05-MAR-2003; 2003WO-SE000372.
 XX
 PR 05-MAR-2002; 2002US-0361333P.
 PR 27-NOV-2002; 2002US-0429307P.
 XX
 PA (ANGI-) ANGIOGENETICS SWEDEN AB.
 XX
 PI Pereira T, Poellinger L, Hellstroem M;
 XX WPI; 2003-712876/67.
 XX
 PT New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.
 XX
 PS Example 6; Fig 19; 96pp; English.
 XX
 CC The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. Sequences AAO23500-509represent mutant fragments
 CC within the N-TAD region of a murine HIF-1 alpha protein
 XX
 SQ Sequence 29 AA;
 Query Match 82.9%; Score 34; DB 7; Length 29;
 Best Local Similarity 87.5%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 DB 15 MLAPYIPM 22
 RESULT 40
 AAO23471
 ID AAO23471 standard; peptide; 29 AA.
 XX
 AC AAO23471;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Murine HIF-1alpha protein fragment.
 XX
 KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy.
 XX
 OS Mus sp.
 XX

XX PT New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing

CC The invention relates to a hypoxia-inducible factor (HIF)-1-alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1-alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1-alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. Sequences AAO23500-509 represent mutant fragments
 CC within the N-TAD region of a murine HIF-1 alpha protein

XX Sequence 29 AA;

Query Match 82.9%; Score 34; DB 7; Length 29;
 Best Local Similarity 87.5%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
 DB 15 MLAPYIPM 22

RESULT 45

AAO23474
 ID AAO23474 standard; peptide; 29 AA.

XX AAO23474;

DT 12-FEB-2004 (first entry)

DE Murine HIF-1-alpha protein mutant fragment.

XX HIF-1-alpha; hypoxia-inducible factor -1-alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW mutant.

OS Mus sp.

PN WO2003074560-A2.

XX 12-SEP-2003.

PD 05-MAR-2003; 2003WO-SE000372.

PF 05-MAR-2002; 2002US-0361333P.

XX 27-NOV-2002; 2002US-0429307P.

XX (ANGI-) ANGIOGENETICS SWEDEN AB.

PA Pereira T, Poellinger L, Hellstroem M;

XX WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1-alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1-alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.

XX Claim 39; Page 90; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1-alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1-alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1-alpha underexpression in a

CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. The present sequence represents a specific example
 CC of a murine HIF-1 alpha mutant fragment used for treatment for hypoxic-
 CC related conditions

XX Sequence 29 AA;

Query Match 82.9%; Score 34; DB 7; Length 29;
 Best Local Similarity 87.5%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
 DB 15 MLAPYIPM 22

RESULT 46

AAO23477
 ID AAO23477 standard; peptide; 29 AA.

XX AAO23477;

DT 12-FEB-2004 (first entry)

DE Murine HIF-1-alpha protein mutant fragment.

XX HIF-1-alpha; hypoxia-inducible factor -1-alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW mutant.

OS Mus sp.

PN WO2003074560-A2.

XX 12-SEP-2003.

PD 05-MAR-2003; 2003WO-SE000372.

PF 05-MAR-2002; 2002US-0361333P.

XX 27-NOV-2002; 2002US-0429307P.

XX (ANGI-) ANGIOGENETICS SWEDEN AB.

PA Pereira T, Poellinger L, Hellstroem M;

XX WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1-alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1-alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.

XX Claim 39; Page 91; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1-alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1-alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1-alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing vascular

CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. The present sequence represents a specific example
 CC of a murine HIF-1 alpha mutant fragment used for treatment for hypoxic-
 CC related conditions

XX Sequence 29 AA;

Query Match 82.9%; Score 34; DB 7; Length 29;

Best Local Similarity 87.5%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLAPTIPM 8

Db 15 MLAPVIMP 22

RESULT 47

AAO23481
 ID AAO23481 standard; peptide; 29 AA.

XX AAO23481;

XX 12-FEB-2004 (first entry)

DE Murine HIF-1alpha protein mutant fragment.

KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW mutant.

XX Mus sp.

PN WO2003074560-A2.

XX 12-SEP-2003.

XX 05-MAR-2003; 2003WO-SE000372.

PR 05-MAR-2002; 2002US-0361333P.

PR 27-NOV-2002; 2002US-0429307P.

XX (ANGI-) ANGIOGENETICS SWEDEN AB.

PI Pereira T, Poellinger L, Hellstrom M;

XX WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.

XX Claim 39; Page 93; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing a hypoxic
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. The present sequence represents a specific example
 CC of a murine HIF-1 alpha mutant fragment used for treatment for hypoxic-

CC related conditions

XX Sequence 29 AA;

Query Match 82.9%; Score 34; DB 7; Length 29;

Best Local Similarity 87.5%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLAPTIPM 8

Db 15 MLAPVIMP 22

RESULT 48

AAO23480
 ID AAO23480 standard; peptide; 29 AA.

XX AAO23480;

XX 12-FEB-2004 (first entry)

DE Murine HIF-1alpha protein mutant fragment.

KW HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
 KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
 KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
 KW mutant.

XX Mus sp.

PN WO2003074560-A2.

XX 12-SEP-2003.

XX 05-MAR-2003; 2003WO-SE000372.

PR 05-MAR-2002; 2002US-0361333P.

PR 27-NOV-2002; 2002US-0429307P.

XX (ANGI-) ANGIOGENETICS SWEDEN AB.

PI Pereira T, Poellinger L, Hellstrom M;

XX WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
 PT angiogenesis, or treating a condition associated with HIF-1alpha
 PT underexpression in a cell, a group of cells, or an organism, e.g.
 PT ischemia or inflammation.

XX Claim 39; Page 93; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
 CC that has (a) an altered transactivation capacity and improved stability
 CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
 CC pharmaceutical composition are useful for increasing angiogenesis,
 CC interfering with a normal response to reoxygenation following hypoxia, or
 CC treating a condition associated with HIF-1alpha underexpression in a
 CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
 CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
 CC stroke. The proteins and pharmaceutical compositions are also useful for
 CC mimicking the hypoxic response or artificially inducing a hypoxic
 CC response in a cell, group of cells, or organism, inducing a hypoxic
 CC formation or vascular development in a cell or a group of cells,
 CC increasing angiogenic activity in a cell, or influencing erythropoietin
 CC production, metabolism, or an inflammatory response in a cell, a group of
 CC cells, or an organism. The present sequence represents a specific example
 CC of a murine HIF-1 alpha mutant fragment used for treatment for hypoxic-
 CC related conditions

XX Sequence 29 AA;

Query Match

82.9%; Score 34; DB 7; Length 29;

Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
|||||
Db 15 MLAPYIPM 22

RESULT 49

AAO23499
ID AAO23499 standard; peptide; 29 AA.

XX AAO23499;

XX 12-FEB-2004 (first entry)

XX Murine HIF-1alpha protein N-TAD region fragment (residues 546-573).

XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW N-TAD.

XX Mus sp.

XX WO2003074560-A2.

XX 12-SEP-2003.

XX 05-MAR-2003; 2003WO-SE000372.

XX 05-MAR-2002; 2002US-0361333P.

XX 27-NOV-2002; 2002US-0429307P.

XX (ANGI-) ANGIOGENETICS SWEDEN AB.

XX Pereira T, Poellinger L, Hellstroem M;

XX WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.

XX Example 6; Fig 19; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1alpha underexpression in a
CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell or a group of cells,
CC increasing angiogenetic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. The present sequence represents a N-TAD region of
CC a murine HIF-1 alpha protein

XX Sequence 29 AA;

Query Match 82.9%; Score 34; DB 7; Length 29;

Best Local Similarity 87.5%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
|||||
Db 15 MLAPYIPM 22

RESULT 50

AAO23500
ID AAO23500 standard; peptide; 29 AA.

XX AAO23500;

XX 12-FEB-2004 (first entry)

XX Murine HIF-1alpha protein N-TAD region mutant fragment 556/558 LL-AA.

XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
KW ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
KW antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
KW N-TAD; mutant.

XX Mus sp.

XX WO2003074560-A2.

XX 12-SEP-2003.

XX 05-MAR-2003; 2003WO-SE000372.

XX 05-MAR-2002; 2002US-0361333P.

XX 27-NOV-2002; 2002US-0429307P.

XX (ANGI-) ANGIOGENETICS SWEDEN AB.

XX Pereira T, Poellinger L, Hellstroem M;

XX WPI; 2003-712876/67.

XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.

XX Example 6; Fig 19; 96pp; English.

XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1alpha underexpression in a
CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell or a group of cells,
CC increasing angiogenetic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. Sequences AAO23500-509 represent mutant fragments
CC within the N-TAD region of a murine HIF-1 alpha protein

XX Sequence 29 AA;

Query Match 82.9%; Score 34; DB 7; Length 29;

Best Local Similarity 87.5%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
|||||
Db 15 MLAPYIPM 22

Search completed: February 9, 2005, 06:05:40

Job time : 53.8246 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 05:55:37 ; Search time 21.193 Seconds
(without alignments)
193.301 Million cell updates/sec

Title: US-10-032-361-4

Perfect score: 41

Sequence: 1 MLAPTPM 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 65 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	298	2 Q6N0X7	Q6N0X7 rhodopsin
2	38	92.7	320	2 Q87LN5	Q87LN5 vibrio para
3	38	92.7	321	2 Q9KPA8	Q9KPA8 vibrio chol
4	37	90.2	588	2 Q6N3P7	Q6N3P7 rhodopsin
5	36	87.8	222	2 Q88VW1	Q88VW1 lactobacill
6	35	85.4	246	2 Q7NSL6	Q7NSL6 chromobacte
7	35	85.4	462	2 Q9A6Q6	Q9A6Q6 caulobacte
8	35	85.4	835	2 Q6K317	Q6K317 oryza sativ
9	34	82.9	120	2 Q7VHC6	Q7VHC6 helicobacte
10	34	82.9	766	1 H1FA_ONCMY	Q988W2 oncorhynch
11	34	82.9	774	2 Q6STN7	Q6STN7 ctenopharyn
12	34	82.9	777	2 Q6EH14	Q6EH14 brachydanio
13	34	82.9	786	2 Q6SLL1	Q6SLL1 canis fami
14	34	82.9	802	2 Q6PI54	Q6PI54 xenopus lae
15	34	82.9	805	1 H1FA_XENLA	Q918A9 xenopus lae
16	34	82.9	811	1 H1FA_CHICK	Q9YIB9 gallus gall
17	34	82.9	819	2 Q7YSE5	Q7YSE5 cryptotag
18	34	82.9	821	2 Q4F454	Q4F454 spermophilu
19	34	82.9	823	1 H1FA_BOVIN	Q9XTA5 bos taurus
20	34	82.9	823	2 Q6IV47	Q6IV47 bos mutus g
21	34	82.9	824	2 Q6H8T3	Q6H8T3 spalia juda
22	34	82.9	825	1 H1FA_RAT	Q35800 rattus norv
23	34	82.9	826	1 H1FA_HUMAN	Q16665 homo sapien
24	34	82.9	836	1 H1FA_MOUSE	Q61221 mus musculu
25	33	80.5	71	1 Y4UG_RHIN	P55671 rhizobium s
26	33	80.5	219	2 Q8MUT9	Q8MUT9 aplysia cal
27	33	80.5	267	1 Y4WMP5	Q7VMP5 haemophilus
28	33	80.5	377	1 Y4WDP_RHIN	P55682 rhizobium s
29	33	80.5	553	2 Q9ATN6	Q9ATN6 capsicum an
30	33	80.5	554	2 Q6D3T4	Q6D3T4 erwinia car
31	33	80.5	1604	2 Q7XQ14	Q7XQ14 oryza sativ

32	32	78.0	145	2	Q826Z2	streptomyce
33	32	78.0	191	2	Q7Y4Y5	bacterioph
34	32	78.0	196	1	VG53_BPT4	P15011 bacterioph
35	32	78.0	243	2	Q8PR29	xanthomonas
36	32	78.0	250	2	Q92LJ9	rhizobium m
37	32	78.0	277	2	Q9M3C6	arabidopsis
38	32	78.0	312	1	PUR3_VIGUN	P52423 vigna ungu
39	32	78.0	335	2	Q9NKD1	dracopis
40	32	78.0	343	2	Q8WE47	draco blaf
41	32	78.0	375	2	Q7QAJ5	anopheles g
42	32	78.0	376	2	Q9MSF3	nepenthes p
43	32	78.0	386	2	Q98H39	rhizobium l
44	32	78.0	401	2	Q9RY74	deinococcus
45	32	78.0	435	2	Q8Y1D4	raistonia s
46	32	78.0	447	2	Q9XEV9	nicotiana t
47	32	78.0	471	2	Q8AVC6	kenopus lae
48	32	78.0	484	2	Q6DJR1	kenopus tro
49	32	78.0	504	2	Q95GF8	nepenthes p
50	32	78.0	504	2	Q6FBU5	acinetobact
51	32	78.0	510	2	Q7YKQ6	byblis giga
52	32	78.0	513	2	Q8MA72	byblis lini
53	32	78.0	611	2	Q6BHW1	debaromyce
54	32	78.0	666	1	NOD_DROME	P13105 drosophila
55	32	78.0	710	2	Q9SX53	arabidopsis
56	32	78.0	723	2	Q07834	pseudomonas
57	32	78.0	741	2	Q8TUS6	methanopyru
58	32	78.0	766	2	Q7R4C0	giardia lam
59	32	78.0	779	2	Q91NU0	arabidopsis
60	32	78.0	821	2	Q8GHR2	pseudomonas
61	32	78.0	907	2	Q8S1S4	oryza sativ
62	32	78.0	931	1	MANB_ASFNG	aspergillus
63	32	78.0	942	2	Q7UGI4	rhodopirell
64	32	78.0	1089	1	NML8_MYCTU	mycobacteri
65	32	78.0	1089	2	Q7TVL0	mycobacteri

ALIGNMENTS

RESULT 1

ID	Q6N0X7	PRELIMINARY;	PRT;	298 AA.
AC	Q6N0X7;			
DT	05-JUL-2004 (Tremblrel. 27, Created)			
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)			
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)			
DE	Short-chain dehydrogenase (EC 1.1.1.100).			
GN	Name=fixr2; OrderedLocusNames=RPA4633;			
OS	Rhodopseudomonas palustris.			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OX	NCBI_TaxID=1076;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CGA009 / ATCC BAA-98;			
RX	PubMed=14704707; DOI=10.1038/nbt923;			
RA	Larimer F.W., Chain P., Hauser L., Lamerdin J.B., Malfatti S., Do L.,			
RA	Land M.L., Palletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,			
RA	Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,			
RA	Harrison F.H., Gibson J., Harwood C.S.;			
RT	"Complete genome sequence of the metabolically versatile			
RL	photosynthetic bacterium Rhodopseudomonas palustris.";			
CC	Nat. Biotechnol. 22:55-61(2004).			
CC	-1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases			
CC	(SDR) family.			
DR	EMBL; BX572607; CAE30073.1; -.			
DR	HSSP; P25529; 1AHI.			
DR	GO; GO:0004316; F:3-oxoacyl-[acyl-carrier protein] reductase . . . ; IEA.			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	InterPro; IPR002198; Adh_short.			
DR	InterPro; IPR002347; Adh_short_C2.			
DR	Pfam; PF00106; adh_short; 1.			

```

DR PRINTS: PR00081; GDRDH.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 298 AA; 31846 MW; 0FCB8D367A162C4 CRC64;

Query Match 100.0%; Score 41; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
Db 254 MLAPTIPM 261

RESULT 2
Q87LN5 PRELIMINARY; PRT; 320 AA.
AC Q87LN5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sigma-E factor regulatory protein RseB.
GN OrderedLocNames=VP2576;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=KMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749 (2003).
DR EMBL; AF005082; BAC60839.1; -.
DR InterPro; IPR005588; MucB_ReeB.
DR Pfam; PF03888; MucB_ReeB; 1.
KW Complete proteome.
SQ SEQUENCE 320 AA; 36042 MW; F8C07A27C63BF2F9 CRC64;

Query Match 92.7%; Score 38; DB 2; Length 320;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
Db 106 MVAPTIPM 113

RESULT 3
Q9KPA8 PRELIMINARY; PRT; 321 AA.
AC Q9KPA8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sigma-E factor regulatory protein RseB.
GN OrderedLocNames=VC2465;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,

```

```

RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483 (2000).
DR EMBL; AE004316; AAF95607.1; -.
DR FIR; F82073; F82073.
DR TIGR; VC2465; -.
DR InterPro; IPR005588; MucB_ReeB.
DR Pfam; PF03888; MucB_ReeB; 1.
KW Complete proteome.
SQ SEQUENCE 321 AA; 36165 MW; A0E94994F41986E1 CRC64;

Query Match 92.7%; Score 38; DB 2; Length 321;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
Db 106 MVAPTIPM 113

RESULT 4
Q6N3P7 PRELIMINARY; PRT; 588 AA.
AC Q6N3P7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative maltotriose trehalose trehalohydrolase (EC 2.4.1.18).
GN OrderedLocNames=RPA3646;
OS Rhodospseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres Y. Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61 (2004).
DR EMBL; BX572604; CAE29087.1; -.
DR GO; GO:0003844; F:1,4-alpha-glucan branching enzyme activity; IEA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_ami cat.
DR InterPro; IPR006589; Alp_ami cat sub.
DR InterPro; IPR004193; Glyco_hydro_13N.
DR Pfam; PF00128; Alpha-amylase; 1.
DR Pfam; PF02922; Isoamylase_N; 1.
DR SMART; SM00642; Amy; 1.
KW Complete proteome.
SQ SEQUENCE 588 AA; 65631 MW; 051C0645E17CD0CF CRC64;

Query Match 90.2%; Score 37; DB 2; Length 588;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
Db 417 LLAPTIPM 424

RESULT 5
Q89VW1

```

```

ID Q88VV1 PRELIMINARY; PRT; 222 AA.
AC Q88VV1;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Phosphoglycolate phosphatase (EC 3.1.3.18).
GN Name=gph2; OrderedLocusNames=gph1932;
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
ON NCBI_TaxID=1590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCF81;
RX MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RT "Complete genome sequence of Lactobacillus plantarum WCF81.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
DR EMBL; AL935257; CAD64320.1; -
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008967; F:phosphoglycolate phosphatase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005834; Dehal like_hydro.
DR Pfam; PF00702; Hydrolase; I.
KW Complete proteome.
SQ SEQUENCE 222 AA; 24559 MW; F5343EC7A91E11F6 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIPM 8
DB 13 LAPTIPM 19

RESULT 6
Q7NSL6 PRELIMINARY; PRT; 246 AA.
AC Q7NSL6;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Probable oxidoreductase.
GN OrderedLocusNames=CV3406;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
ON NCBI_TaxID=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.18332124100;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.,
RA Azeiteiro-Filho S., Azevedo V., Baptista A.J., Batais L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
RA Bordignon J., Brigo M.M., Brito C.A., Brocchi M., Burity H.A.,
RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furian L.R.,
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Grattapaglia D., Grissard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manlio G.P., Maranhao A.Q., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,

```

```

RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seanez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
RA Vettore A., Wassem R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
DR EMBL; AB016922; AAQ61070.1; -
DR HSSP; P19992; 1HDC.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; ADH_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRDH.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 246 AA; 25691 MW; 40058ABA52395E22 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 246;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIPM 8
DB 205 LAPTIPM 211

RESULT 7
Q9A6Q6 PRELIMINARY; PRT; 462 AA.
AC Q9A6Q6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein CC2028.
GN OrderedLocusNames=CC2028;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteriaceae; Caulobacter.
ON NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Neilson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Neilson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Sait J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermlaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005876; AAK24003.1; -
DR PIR; G87500; G87500.
DR TIGR; CC2028; -
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001511; Peptidase_S41.
DR Pfam; PF03572; Peptidase_S41; I.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 462 AA; 49123 MW; A4F08680B10DFBB2 CRC64;

```

Query Match 85.4%; Score 35; DB 2; Length 462;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIMP 8
 |||||
 Db 423 LAPIVPM 429

RESULT 8
 Q6K317 PRELIMINARY; PRT; 835 AA.
 AC Q6K317;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Putative vegetative cell wall protein gp1precursor.
 GN Name=OSJNB0066C12.14;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
 RT clone:OSJNB0066C12.14";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005738; BAD23553.1; -;
 DR InterPro; IPR006162; Ppantne_S.
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; UNKNOWN 1.
 SQ SEQUENCE 835 AA; 86306 MW; CA76DDA761E8AE41 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 835;
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIMP 8
 |||||
 Db 524 MLAPPIPM 531

RESULT 9
 Q7VHC6 PRELIMINARY; PRT; 120 AA.
 AC Q7VHC6;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=HH1041;
 OS Helicobacter hepaticus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=32025;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 51449 / 3B1;
 RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
 RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
 RA Bell M., Droege M., Farman B., Fischer H.-P., Ge Z., Hoerster A.,
 RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
 RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
 RT "The complete genome sequence of the carcinogenic bacterium
 RT Helicobacter hepaticus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
 DR EMBL; AF017147; AAP77638.1; -;
 KW Complete proteome.
 SQ SEQUENCE 120 AA; 13629 MW; DA6F756A8992A9B8 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 120;
 Best Local Similarity 85.7%; Pred. No. 45;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 LAPTIMP 8
 |||||
 Db 103 LAPTEPM 109

RESULT 10
 HIFA_ONCMY STANDARD; PRT; 766 AA.
 AC Q98SW2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).
 GN Name=HIF1A;
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21282949; PubMed=11278461; DOI=10.1074/jbc.M009057200;
 RA Soitamo A.J., Rabergh C.M.I., Gassmann M., Sistonen L., Nikimaa M.;
 RT "Characterization of a hypoxia-inducible factor (HIF-1 alpha) from
 RT rainbow trout: accumulation of protein occurs at normal venous oxygen
 RT tension.";
 RL J. Biol. Chem. 276:19699-19705(2001).
 CC -1- FUNCTION: Functions as a master transcriptional regulator of the
 CC adaptive response to hypoxia. Binds to core DNA sequence 5'-
 CC [AG]CGTG-3' within the hypoxia response element (HRE) of target
 CC gene promoters. Activation requires recruitment of transcriptional
 CC coactivators (By similarity).
 CC -1- SUBUNIT: Efficient DNA binding requires heterodimerization of an
 CC alpha and a beta/ARNT subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear
 CC translocation in response to hypoxia (By similarity).
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF304864; AAK30364.1; -;
 DR HSSP; Q16665; 1H2K.
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR001067; Nuc_translocat.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PRINTS; PR00785; NCTRNSLOCATR.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 1.
 DR PROSITE; PS50888; HLH; 1.
 DR PROSITE; PS50112; PAS; 2.
 DR Activator; DNA-binding; Nuclear protein; Repeat;
 KW Transcription regulation.
 FT DOMAIN 718 721 Nuclear localization signal (Potential).
 FT DNA BIND 17 30 Basic motif.
 FT DOMAIN 31 71 Helix-loop-helix motif.
 FT DOMAIN 82 159 PAS 1.
 FT DOMAIN 230 300 PAS 2.
 FT DOMAIN 304 347 PAC.

```

FT DOMAIN 368 373 Poly-Glu.
SQ SEQUENCE 766 AA; 85049 MW; FC25A4984104DA5B CRC64;

Query Match 82.9%; Score 34; DB 1; Length 766;
Best Local Similarity 87.5%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPITPM 8
Db 556 MLAPYIPM 563

RESULT 11
Q6STN7 PRELIMINARY; PRT; 774 AA.
AC Q6STN7;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DE Hypoxia-inducible factor-1alpha.
GN Name=hif-1alpha;
OS Ctenopharyngodon idella (Grass carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Ctenopharyngodon.
OX NCBI_TaxID=7959;
RN [1]
RP SEQUENCE FROM N.A.
RA Law S.H.W., Wu R.S.S., Mok H.O.L., Yu R.M.K., Ng P.K.S., Kong R.Y.C.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH basic.
DR InterPro; IPR001321; HypoxindF1A.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00785; PAC; 1.
DR PRINTS; PR01080; HYPOXIAF1A.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS50888; HLH; 1.
DR PROSITE; PS50112; PAS; 2.
SQ SEQUENCE 774 AA; 85741 MW; 15BD0B2F90C611C6 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 774;
Best Local Similarity 87.5%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPITPM 8
Db 552 MLAPYIPM 559

RESULT 12
Q6EH14 PRELIMINARY; PRT; 777 AA.
AC Q6EH14;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE Hypoxia-inducible factor 1 alpha.
GN Name=hif1a;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

```

```

OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang C.-R., Tsai Y.-C., Young H.-W., Wang W.-D., Hu C.-H.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AY326951; AAQ91619.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH basic.
DR InterPro; IPR001321; HypoxindF1A.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00785; PAC; 1.
DR PRINTS; PR01080; HYPOXIAF1A.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS50888; HLH; 1.
DR PROSITE; PS50112; PAS; 2.
SQ SEQUENCE 777 AA; 85897 MW; A9AFE2D2732C7A33 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 777;
Best Local Similarity 87.5%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPITPM 8
Db 554 MLAPYIPM 561

RESULT 13
Q6SLL1 PRELIMINARY; PRT; 786 AA.
AC Q6SLL1;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DE Hypoxia-inducible factor 1 alpha subunit (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Spee B., Penning L.C., Rothuizen J.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY455802; AAR19225.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; P:signal transducer activity; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001092; HLH basic.
DR InterPro; IPR001321; HypoxindF1A.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR01080; HYPOXIAF1A.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS50112; PAS; 2.
FT NON_TER 1 786 786

```

```

SQ SEQUENCE 786 AA; 88015 MW; C37A27C25C343CDC CRC64;
Query Match 82.9%; Score 34; DB 2; Length 786;
Best Local Similarity 87.5%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPITPM 8
DB 541 MLAPVPM 548

RESULT 14
Q6PI54 PRELIMINARY; PRT; 802 AA.
AC Q6PI54;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DE Hfla-prov protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;
RA Klein S., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (BHLH) domain.
DR EMBL; BC043769; AA043769.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH basic.
DR InterPro; IPR001321; Hypoxindf1A.
DR InterPro; IPR001610; PAC.
DR PROSITE; PS00091; PAS; 2.
DR PROSITE; PS50888; HLH; FALSE_NEG.

RESULT 15
HIFA XENLA STANDARD; PRT; 805 AA.
AC Q918X9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).
GN Name=HIF1A;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Kietzmann T.;
RT "Cloning and expression of the Xenopus laevis hypoxia inducible factor
RL 1 alpha homologue."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Functions as a master transcriptional regulator of the
CC adaptive response to hypoxia. Binds to core DNA sequence 5'-
CC [AG]CGTG-3' within the hypoxia response element (HRE) of target
CC gene promoters. Activation requires recruitment of transcriptional
CC coactivators (by similarity).
CC -1- SUBUNIT: Efficient DNA binding requires heterodimerization of an
CC alpha and a beta/ARNT subunit (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear
CC translocation in response to hypoxia (by similarity).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (BHLH) domain.
CC -1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR HSSP; Q16665; 1H2K.
DR InterPro; IPR001092; HLH basic.
DR InterPro; IPR001321; Hypoxindf1A.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR Pfam; PF00785; PAS; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR01080; HYPOXIAF1A.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS50888; HLH; FALSE_NEG.

```


DR PROSITE; PS50112; PAS; 2.
 KW Activator; DNA-binding; Nuclear protein; Repeat;
 Transcription regulation.
 FT DNA_BIND 17 30 Basic motif.
 FT DOMAIN 31 71 Helix-loop-helix motif.
 FT DOMAIN 85 157 PAS 1.
 FT DOMAIN 229 300 PAS 2.
 FT DOMAIN 303 346 PAC.
 SQ SEQUENCE 805 AA; 90964 MW; BABFA0BD6B44FF3B CRC64;

Query Match 82.9%; Score 34; DB 1; Length 805;
 Best Local Similarity 87.5%; Pred. No. 3.2e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIMP 8
 |||||
 Db 557 MLAPYIPM 564

RESULT 16
 HIFA CHICK
 ID HIFA CHICK STANDARD; PRT; 811 AA.
 AC Q9VIB9.
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).
 GN Name=HIF1A;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RA Takahashi T.;
 RT "Molecular cloning and expression of an avian cDNA for hypoxia-inducible factor-1 alpha in embryonic ventricular myocytes.";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Functions as a master transcriptional regulator of the adaptive response to hypoxia. Binds to core DNA sequence 5'-[AG]CGG-3' within the hypoxia response element (HRE) of target gene promoters. Activation requires recruitment of transcriptional coactivators (By similarity).
 CC -!- SUBUNIT: Efficient DNA binding requires heterodimerization of an alpha and a beta/ARNT subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear translocation in response to hypoxia (By similarity).
 CC -!- DOMAIN: Contains two independent C-terminal transactivation domains, NTAD and CTAD, which function synergistically. Their transcriptional activity is repressed by an intervening inhibitory domain (ID) (By similarity).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -!- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
 CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; AB013746; BAA34234.2; -.
 DR HSSP; Q16665; 1H2K.
 DR InterPro; IPR001092; HLH basic.
 DR InterPro; IPR001321; HypoxindfIA.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00785; PAC; 1.

DR Pfam; PF00989; PAS; 2.
 DR PRINTS; PRO1080; HYPOXIAIFIA.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS50888; HLH; 1.
 DR PROSITE; PS50112; PAS; 2.
 KW Activator; DNA-binding; Nuclear protein; Repeat;
 Transcription regulation.
 FT DNA_BIND 17 30 Basic motif.
 FT DOMAIN 31 71 Helix-loop-helix motif.
 FT DOMAIN 80 157 PAS 1.
 FT DOMAIN 228 298 PAS 2.
 FT DOMAIN 302 345 PAC.
 FT DOMAIN 401 587 ODD.
 FT DOMAIN 529 573 NTAD.
 FT DOMAIN 576 785 ID.
 FT DOMAIN 703 706 Nuclear localization signal (Potential).
 FT DOMAIN 718 721 Nuclear localization signal (Potential).
 FT DOMAIN 771 811 CTAD.
 FT DOMAIN 583 588 Poly-Ser.
 SQ SEQUENCE 811 AA; 90542 MW; D14CD9FC98F064CB CRC64;

Query Match 82.9%; Score 34; DB 1; Length 811;
 Best Local Similarity 87.5%; Pred. No. 3.2e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIMP 8
 |||||
 Db 559 MLAPYIPM 566

RESULT 17
 Q7YSE5
 ID Q7YSE5 PRELIMINARY; PRT; 819 AA.
 AC Q7YSE5;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypoxia inducible factor 1 alpha subunit.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Clausen I., Kietz S., Fischer B.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 DR EMBL; AY273790; AAP43517.1; -.
 DR HSSP; Q16665; 1H2K.
 DR GO; GO:0004871; P:signal transducer activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007185; P:signal transduction; IEA.
 DR InterPro; IPR001092; HLH basic.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 2.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS50888; HLH; 1.
 DR PROSITE; PS50112; PAS; 2.
 SQ SEQUENCE 819 AA; 91284 MW; E11B4PBF7D4F6C7C CRC64;

Query Match 82.9%; Score 34; DB 2; Length 819;
 Best Local Similarity 87.5%; Pred. No. 3.3e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIMP 8
 |||||

[illegible]

33

```

FT MOD_RES 334 334 S-nitrosocysteine (Potential).
FT MOD_RES 337 337 S-nitrosocysteine (Potential).
FT MOD_RES 359 359 S-nitrosocysteine (Potential).
FT MOD_RES 402 402 Hydroxyproline (By similarity).
FT MOD_RES 520 520 S-nitrosocysteine (Potential).
FT MOD_RES 532 532 N6-acetyllysine (By similarity).
FT MOD_RES 564 564 Hydroxyproline (By similarity).
FT MOD_RES 755 755 S-nitrosocysteine (Potential).
FT MOD_RES 777 777 S-nitrosocysteine (Potential).
FT MOD_RES 797 797 S-nitrosocysteine (Potential).
FT MOD_RES 800 800 3-hydroxyasparagine (By similarity).
SQ SEQUENCE 823 AA; 92127 MW; 12674E467A61B1A1 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 823;
Best Local Similarity 87.5%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPYIPM 8
Db 561 MLAPYIPM 568

RESULT 20
ID Q6IV47 PRELIMINARY; PRT; 823 AA.
AC Q6IV47;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypoxia inducible factor-1a.
GN Name=HIF-1a;
OS Bos mutus grunniens (Yak).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OC NCBI_TaxID=30521;
RN [1]
RP SEQUENCE FROM N.A.
RA Dolt K.S., Qadar Pasha M.A.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AY621118; AAT39520.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001321; HypoxindFIA.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAC; 1.
DR PRINTS; PR01080; HYPOXIAFIA.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAC; 2.
DR PROSITE; PS50888; HLH; 1.
DR PROSITE; PS50112; PAS; 2.
SQ SEQUENCE 823 AA; 92128 MW; A6388B54F8A15705 CRC64;

Query Match 82.9%; Score 34; DB 2; Length 823;
Best Local Similarity 87.5%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPYIPM 8
Db 561 MLAPYIPM 568

RESULT 21
Q6H8T3

```

```

ID Q6H8T3 PRELIMINARY; PRT; 824 AA.
AC Q6H8T3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypoxia inducible factor 1 alpha.
GN Name=hif-1a;
OS Spalax judaei (Blind subterranean mole rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Spalacinae;
OC Spalax.
OC NCBI_TaxID=134510;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Shams I., Avioli A., Nevo E.;
RT "Hypoxic stress tolerance of the blind subterranean mole rat: Expression of erythropoietin and hypoxia-inducible factor-1a.";
RL Nucleic Acids Res. 0:0-0(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
EX PubMed=15210955; DOI=10.1073/pnas.0403540101;
RA Shams I., Avioli A., Eviatar N.;
RT "Hypoxic stress tolerance of the blind subterranean mole rat: expression of erythropoietin and hypoxia-inducible factor 1 alpha.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9698-9703(2004).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AJ715791; CAG29396.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001321; HypoxindFIA.
DR InterPro; IPR001610; PAC.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAC; 2.
DR PRINTS; PR01080; HYPOXIAFIA.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAC; 2.
DR PROSITE; PS50888; HLH; 1.
DR PROSITE; PS50112; PAS; 2.
SQ SEQUENCE 824 AA; 92161 MW; 33A1DDC3593CBFFF CRC64;

Query Match 82.9%; Score 34; DB 2; Length 824;
Best Local Similarity 87.5%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPYIPM 8
Db 561 MLAPYIPM 568

RESULT 22
HIFA RAT
ID HIFA RAT STANDARD; PRT; 825 AA.
AC Q35800; Q9WTU9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha).
GN Name=Hif1a;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=Wistar; TISSUE=Hepatocytes;
 RX MEDLINE=21134367; PubMed=11237857; DOI=10.1042/0264-6021.3540531;
 RA Kietzmann T., Cornesse Y., Brechtel K., Modaresi S., Jungermann K.;
 RT "Perivascular expression of the mRNA of the three hypoxia-inducible
 factor a-subunits HIF-1a, HIF2a and HIF3a in rat liver";
 RL Biochem. J. 354:531-537(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
 RX MEDLINE=21417706; PubMed=11526200;
 RA Zou A.-P., Yang Z.-Z., Li P.-L., Cowley A.W. Jr.;
 RT "Oxygen-dependent expression of hypoxia-inducible factor-1alpha in
 renal medullary cells of rats";
 RL Physiol. Genomics 6:159-168(2001).
 CC -1- FUNCTION: Functions as a master transcriptional regulator of the
 CC adaptive response to hypoxia. Under hypoxic conditions activates
 CC the transcription of over 40 genes, including, erythropoietin,
 CC glucose transporters, glycolytic enzymes, vascular endothelial
 CC growth factor, and other genes whose protein products increase
 CC oxygen delivery or facilitate metabolic adaptation to hypoxia.
 CC Plays an essential role in embryonic vascularization, tumor
 CC angiogenesis and pathophysiology of ischemic disease. Binds to
 CC core DNA sequence 5'-[AG]GTG-3' within the hypoxia response
 CC element (HRE) of target gene promoters. Activation requires
 CC recruitment of transcriptional coactivators such as CREBBP and
 CC EP300. Activity is enhanced by interaction with both, NCOAL or
 CC NCOA2. Interaction with redox regulatory protein APEX seems to
 CC activate CTAD and potentiates activation by NCOAL and CREBBP (By
 CC similarity).
 CC -1- SUBUNIT: Efficient DNA binding requires heterodimerization of an
 CC alpha and a beta/ARNT subunit. Binds to the TAZ-type 1 domains of
 CC CREBBP and EP300. Interacts with NCOAL, NCOA2, APEX and HSP90 (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear
 CC translocation in response to hypoxia (By similarity).
 CC -1- TISSUE SPECIFICITY: Expressed in the kidney, higher expression is
 CC seen in the renal medulla than in the cortex. Expressed also in
 CC the perivascular zone of the liver.
 CC -1- DOMAIN: Contains two independent C-terminal transactivation
 CC domains, NTA and CTAD, which function synergistically. Their
 CC transcriptional activity is repressed by an intervening inhibitory
 CC domain (ID) (By similarity).
 CC -1- PTM: In normoxia, is hydroxylated on Pro-402 and Pro-563 in the
 CC oxygen-dependent degradation domain (ODD) by EGLN1/PHD1 and
 CC EGLN2/PHD2. EGLN3/PHD3 has also been shown to hydroxylate Pro-563.
 CC The hydroxylated prolines promote interaction with VHL, initiating
 CC rapid ubiquitination and subsequent proteasomal degradation. Under
 CC hypoxia, proline hydroxylation is impaired and ubiquitination is
 CC attenuated, resulting in stabilization (By similarity).
 CC -1- PTM: In normoxia, is hydroxylated on Asn-802 by HIF1AN, thus
 CC abrogating interaction with CREBBP and EP300 and preventing
 CC transcriptional activation (By similarity).
 CC -1- PTM: S-nitrosylated. All free thiol groups are subjected to S-
 CC nitrosylation in vitro, however not all thiol groups seem to be
 CC nitrosylated in vivo (By similarity).
 CC -1- PTM: Acetylation of Lys-531 by ARD1 increases interaction with VHL
 CC and stimulates subsequent proteasomal degradation (By similarity).
 CC -1- PTM: Phosphorylation is required for DNA binding (By similarity).
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -1- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
 CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y09507; CAA70701.1; --
 CC EMBL; AF057308; AAD24413.1; --
 CC HSSP; Q16665; 118C.

DR TRANSFAC; T05461; --
 DR InterPro; IPR001092; HLH_basic.
 DR InterPro; IPR001321; HypoxindfIA.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR000014; PAS.
 DR Pfam; PF00010; HLH; 1.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 2.
 DR PRINTS; PRO1080; HYPOXIAFIA.
 DR SMART; SM00353; HLH; 1.
 DR SMART; SM00086; PAC; 1.
 DR SMART; SM00091; PAS; 2.
 DR PROSITE; PS00888; HLH; 1.
 DR PROSITE; PS0112; PAS; 2.
 KW Acetylation; Activator; DNA-binding; Hydroxylation; Nuclear protein;
 KW Phosphorylation; Repeat; S-nitrosylation; Transcription regulation.
 FT DOMAIN 718 721
 FT DNA BIND 17 30
 FT DOMAIN 31 71
 FT DOMAIN 85 158
 FT DOMAIN 228 298
 FT DOMAIN 302 345
 FT DOMAIN 401 602
 FT DOMAIN 530 574
 FT DOMAIN 575 784
 FT DOMAIN 717 720
 FT DOMAIN 785 825
 FT MOD RES 90 90
 FT MOD RES 139 139
 FT MOD RES 173 173
 FT MOD RES 194 194
 FT MOD RES 210 210
 FT MOD RES 219 219
 FT MOD RES 224 224
 FT MOD RES 255 255
 FT MOD RES 334 334
 FT MOD RES 337 337
 FT MOD RES 385 385
 FT MOD RES 402 402
 FT MOD RES 519 519
 FT MOD RES 531 531
 FT MOD RES 563 563
 FT MOD RES 779 779
 FT MOD RES 799 799
 FT MOD RES 802 802
 FT CONFLICT 12 12
 FT CONFLICT 74 74
 FT CONFLICT 96 96
 FT CONFLICT 329 329
 FT CONFLICT 613 619
 FT CONFLICT 708 708
 SQ SEQUENCE 825 AA; 92319 MW; C4109A57F38667E9 CRC64;
 Query Match 82.9%; Score 34; DB 1; Length 825;
 Best Local Similarity 87.5%; Pred. No. 3.3e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 Db 560 MLAPVIPM 567
 RESULT 23
 ID HIFA HUMAN STANDARD; PRT; 826 AA.
 AC Q16665; Q96P79; Q9UPB1;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 25-OCT-2004 (Rel. 45; Last annotation update)
 DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (ARNT
 DE interacting protein) (Member of PAS protein 1)
 GN Name=HIF1A;
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 166-170; 259-289 AND 771-781.
 RX MEDLINE=95296340; PubMed=7539918;
 RA Wang G.-L., Jiang B.-H., Rue E.A., Semenza G.L.;
 RT "Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS
 RT heterodimer regulated by cellular O2 tension.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5510-5514 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hepatoma;
 RX MEDLINE=97236817; PubMed=9079689; DOI=10.1074/jbc.272.13.8581;
 RA Hogenesch J.B., Chan W.K., Jackiw V.H., Brown R.C., Gu Y.-Z.,
 RA Pray-Grant M., Perdev G.H., Bradford C.A.;
 RT "Characterization of a subset of the basic-helix-loop-helix-PAS
 RT superfamily that interacts with components of the dioxin signaling
 RT pathway.";
 RL J. Biol. Chem. 272:8581-8593 (1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=99000835; PubMed=9782081; DOI=10.1006/geno.1998.5416;
 RA Iyer N.V., Leung S.W., Semenza G.L.;
 RT "The human hypoxia-inducible factor 1alpha gene: HIF1A structure and
 RT evolutionary conservation.";
 RL Genomics 52:159-165 (1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Rupert J.L., Hochachka P.W.;
 RT "HIF1A sequence in the Quechua, a high altitude population.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Glial tumor;
 RA Sun B., Zhao H.R., Yu R.T., Ni M.S.H.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Liver;
 RA Tanaka S., Sugimachi K.;
 RT "Hypoxia-inducible factor-1 alpha variant isolated from human liver
 RT tissue.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Choriocarcinoma, and Placenta;
 RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grincham J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [8]
 RP TRANSACTIVATION DOMAINS NTAD AND CTAD
 RX MEDLINE=97382249; PubMed=9235919; DOI=10.1074/jbc.272.31.19253;
 RA Jiang B.H., Zheng J.Z., Leung S.W., Roe R., Semenza G.L.;
 RT "Transactivation and inhibitory domains of hypoxia-inducible factor

RT 1alpha. Modulation of transcriptional activity by oxygen tension.";
 RL J. Biol. Chem. 272:19253-19260 (1997).
 RN [9]
 RP SUBCELLULAR LOCATION, AND MUTAGENESIS OF LYS-719.
 RX MEDLINE=95043864; PubMed=9822602; DOI=10.1093/emboj/17.22.6573;
 RA Kallio P.J., Okamoto K., O'Brien S., Carrero P., Makino Y., Tanaka H.,
 RA Poellinger L.;
 RT "Signal transduction in hypoxic cells: inducible nuclear translocation
 RT and recruitment of the CBP/p300 coactivator by the hypoxia-inducible
 RT factor-1alpha.";
 RL EMBO J. 17:6573-6586 (1998).
 RN [10]
 RP OXYGEN-DEPENDENT DEGRADATION DOMAIN.
 RX MEDLINE=98318598; PubMed=9653127; DOI=10.1073/pnas.95.14.7987;
 RA Huang L.E., Gu J., Schau M., Bunn H.F.;
 RT "Regulation of hypoxia-inducible factor 1alpha is mediated by an O2-
 RT dependent degradation domain via the ubiquitin-proteasome pathway.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:7987-7992 (1998).
 RN [11]
 RP TRANSACTIVATION DOMAINS NTAD AND CTAD, INTERACTION WITH APEX, AND
 RP MUTAGENESIS OF CVS-800.
 RX MEDLINE=99219869; PubMed=10202154; DOI=10.1093/emboj/18.7.1905;
 RA Ema M., Hirota K., Mimura J., Abe H., Yodoi J., Sogawa K.,
 RA Poellinger L., Fujii-Kuriyama Y.;
 RT "Molecular mechanisms of transcription activation by HLF and HIF1alpha
 RT in response to hypoxia: their stabilization and redox signal-induced
 RT interaction with CBP/p300.";
 RL EMBO J. 18:1905-1914 (1999).
 RN [12]
 RP INTERACTION WITH NCOA1, NCOA2 AND APEX.
 RX MEDLINE=20063199; PubMed=10594042;
 RA Carrero P., Okamoto K., Coumalleau P., O'Brien S., Tanaka H.,
 RA Poellinger L.;
 RT "Redox-regulated recruitment of the transcriptional coactivators CREB-
 RT binding protein and SRC-1 to hypoxia-inducible factor 1alpha.";
 RL Mol. Cell. Biol. 20:402-415 (2000).
 RN [13]
 RP MUTAGENESIS OF SER-551 AND THR-552.
 RX MEDLINE=20243767; PubMed=10758161; DOI=10.1073/pnas.080072497;
 RA Sutter C.H., Laughner E., Semenza G.L.;
 RT "Hypoxia-inducible factor 1alpha protein expression is controlled by
 RT oxygen-regulated ubiquitination that is disrupted by deletions and
 RT missense mutations.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4748-4753 (2000).
 RN [14]
 RP UBIQUITINATION.
 RX MEDLINE=21214630; PubMed=11292861; DOI=10.1126/science.1059796;
 RA Jaakkola P., Mole D.R., Tian Y.-M., Wilson M.I., Gielbert J.,
 RA Gaskell S.J., von Kriegsheim A., Hebestreit H.F., Mukherji M.,
 RA Schofield C.J., Maxwell P.H., Pugh C.W., Ratcliffe P.J.;
 RT "Targeting of HIF-1alpha to the von Hippel-Lindau ubiquitylation
 RT complex by O2-regulated prolyl hydroxylation.";
 RL Science 292:468-472 (2001).
 RN [15]
 RP S-NITROSYLATION.
 RX MEDLINE=22448624; PubMed=12560087; DOI=10.1016/S0014-5793(02)03887-5;
 RA Sumbayev V.V., Budde A., Zhou J., Bruene B.;
 RT "HIF-1 alpha protein as a target for S-nitrosation.";
 RL FEBS Lett. 535:106-112 (2003).
 RN [16]
 RP ACETYLATION OF LYS-532.
 RX MEDLINE=22351901; PubMed=12464182; DOI=10.1016/S0092-8674(02)01085-1;
 RA Jeong J.-W., Bae M.-K., Ahn M.-Y., Kim S.-H., Sohn T.-K., Bae M.-H.,
 RA Yoo M.-A., Song E.-J., Lee K.-J., Kim K.-W.;
 RT "Regulation and destabilization of HIF-1alpha by ARD1-mediated
 RT acetylation.";
 RL Cell 111:709-720 (2002).
 RN [17]
 RP HYDROXYLATION OF ASN-803.
 RX MEDLINE=22074910; PubMed=12080085; DOI=10.1101/gad.991402;
 RA Lando D., Peet D.J., Gorman J.J., Whelan D.A., Whitelaw M.L.,
 RA Bruick R.K.;
 RT "FIH-1 is an asparaginyl hydroxylase enzyme that regulates the

RT transcriptional activity of hypoxia-inducible factor."; Genes Dev. 16:1466-1471(2002).
RL [18]
RN HYDROXYLATION OF PRO-402 AND PRO-564.
RX MEDLINE=21558830; PubMed=11598268; DOI=10.1126/science.1066373;
RA Bruick R.K., McKnight S.L.;
RT "A conserved family of prolyl-4-hydroxylases that modify HIF."; Science 294:1337-1340(2001).
RL [19]
RN REVIEW.
RX MEDLINE=20407247; PubMed=10950862;
RA Semenza G.L.;
RT "HIF-1 and human disease: one highly involved factor."; Genes Dev. 14:1983-1991(2000).
RL [20]
RN 3D-STRUCTURE MODELING.
RX MEDLINE=20539371; PubMed=11089639;
RA Michel G., Minet E., Ernest I., Roland I., Durant F., Remacle J., Michiels C.;
RT "A model for the complex between the hypoxia-inducible factor-1 (HIF-1) and its consensus DNA sequence."; J. Biol. Struct. Dyn. 18:169-179(2000).
RL [21]
RN X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 775-826 IN COMPLEX WITH HIF1AN.
RX MEDLINE=22412289; PubMed=12446723; DOI=10.1074/jbc.C200644200;
RA Elkins J.M., Hewitson K.S., McNeill L.A., Seibel J.F., Schofield C.J.;
RT "Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 alpha."; J. Biol. Chem. 278:1802-1806(2003).
RL [22]
RN STRUCTURE BY NMR OF 786-826 IN COMPLEX WITH 302-418 OF EP300.
RX MEDLINE=21957254; PubMed=11959990; DOI=10.1073/pnas.082117899;
RA Freedman S.J., Sun Z.-Y.J., Poy F., Kung A.L., Livingston D.M., Wagner G., Eck M.J.;
RT "Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha."; Proc. Natl. Acad. Sci. U.S.A. 99:5367-5372(2002).
RL [23]
RN STRUCTURE BY NMR OF 776-826 IN COMPLEX WITH 345-439 OF CREBBP.
RX MEDLINE=21957241; PubMed=11959977; DOI=10.1073/pnas.082121399;
RA Dames S.A., Martinez-Yamout M., De Guzman R.N., Dyson H.J., Wright P.E.;
RT "Structural basis for Hif-1 alpha /CBP recognition in the cellular Query Match 82.9%; Score 34; DB 1; Length 826;
Best Local Similarity 87.5%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAPTIM 8
DB 561 MLAPYIPM 568
RESULT 24
HIFA_MOUSE
ID HIFA_MOUSE STANDARD; PRT; 836 AA.
AC Q61221; Q08741; Q08993; Q61664; Q61665; Q8C681; Q8CC19; Q8CCB6;
AC Q8R385; Q9CYA8;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypoxia-inducible factor 1 alpha (HIF-1 alpha) (HIF1 alpha) (ARNT
DE interacting protein).
GN Name:Hifia;
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RX STRAIN=C57BL/6; TISSUE=Hepatocytes;

RX MEDLINE=96355491; PubMed=8702901; DOI=10.1074/jbc.271.35.21262;
RA Li H., Ko H.P., Whitlock J.P. Jr.;
RT "Induction of phosphoglycerate kinase 1 gene expression by hypoxia. Roles of Arnt and Hif1alpha."; J. Biol. Chem. 271:21262-21267(1996).
RL [2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RX STRAIN=129/SVJ;
RC MEDLINE=98034461; PubMed=9368100;
RX Luo G., Gu Y.-Z., Jain S., Chan W.K., Carr K.M., Hogenesch J.B., Bradfield C.A.;
RA "Molecular characterization of the murine Hif-1 alpha locus."; Gene Expr. 6:287-299(1997).
RL [3]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RX STRAIN=129/SV;
RC MEDLINE=97354184; PubMed=9210478;
RX Wenger R.H., Rolfs A., Kvietikova I., Spielmann P., Zimmermann D.R., Cassmann M.;
RA "The mouse gene for hypoxia-inducible factor-1alpha. Genomic organization, expression and characterization of an alternative first exon and 5' flanking sequence."; Eur. J. Biochem. 246:155-165(1997).
RL [4]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RX STRAIN=C57BL/6J; TISSUE=Colon; Embryo, and Skin;
RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nakado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bratt D., Brusci V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Saitanin A., Schneider C., Sample C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."; Nature 420:563-573(2002).
RL [5]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=Breast tumor;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [6]
 RN SEQUENCE OF 13-822 FROM N.A. (ISOFORM 2).
 RC TISSUE=Hepatocytes;
 RX MEDLINE=96254028; PubMed=8660378; DOI=10.1006/bbrc.1996.0845;
 RA Wenger R.H., Rolfs A., Marti H.H., Guenet J.-L., Gassmann M.,
 RT "Nucleotide sequence, chromosomal assignment and mRNA expression of
 RT mouse hypoxia-inducible factor-1 alpha.";
 RL Biochem. Biophys. Res. Commun. 223:54-59 (1996).
 RN [7]
 RN SEQUENCE OF 22-85 FROM N.A.
 RC TISSUE=Hepatocytes;
 RA O'Rourke J.F.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Functions as a master transcriptional regulator of the
 CC adaptive response to hypoxia. Under hypoxic conditions activates
 CC the transcription of over 40 genes, including, erythropoietin,
 CC glucose transporters, glycolytic enzymes, vascular endothelial
 CC growth factor, and other genes whose protein products increase
 CC oxygen delivery or facilitate metabolic adaptation to hypoxia.
 CC Plays an essential role in embryonic vascularization, tumor
 CC angiogenesis and pathophysiology of ischemic disease. Binds to
 CC core DNA sequence 5'-[AG]CGTG-3', within the hypoxia response
 CC element (HRE) of target gene promoters. Activation requires
 CC recruitment of transcriptional coactivators such as CREBBP and
 CC EP300. Activity is enhanced by interaction with both, NCOA1 or
 CC NCOA2. Interaction with redox regulatory protein APEX seems to
 CC activate CTAD and potentiates activation by NCOA1 and CREBBP (By
 CC similarity).
 CC -!- SUBUNIT: Efficient DNA binding requires heterodimerization of an
 CC alpha and a beta/ARNT subunit. Binds to the TAZ-type 1 domains of
 CC CREBBP and EP300. Interacts with NCOA1, NCOA2, APEX and HSP90 (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic in normoxia, nuclear
 CC translocation in response to hypoxia (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q61221-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q61221-2; Sequence=VSP_007739;
 CC -!- TISSUE SPECIFICITY: Ubiquitous.
 CC -!- DOMAIN: Contains two independent C-terminal transactivation
 CC domains, NTAD and CTAD, which function synergistically. Their
 CC transcriptional activity is repressed by an intervening inhibitory
 CC domain (ID) (By similarity).
 CC -!- PTM: In normoxia, is hydroxylated on Pro-402 and Pro-577 in the
 CC oxygen-dependent degradation domain (ODD) by EGLN1/PHD1 and
 CC EGLN2/PHD2. EGLN3/PHD3 has also been shown to hydroxylate Pro-577.
 CC The hydroxylated prolines promote interaction with VHL, initiating
 CC rapid ubiquitination and subsequent proteasomal degradation. Under
 CC hypoxia, proline hydroxylation is impaired and ubiquitination is
 CC attenuated, resulting in stabilization (By similarity).
 CC -!- PTM: In normoxia, is hydroxylated on Asn-813 by HIF1AN, thus
 CC abrogating interaction with CREBBP and EP300 and preventing
 CC transcriptional activation (By similarity).
 CC -!- PTM: S-nitrosylated. All 15 free thiol groups are subjected to S-
 CC nitrosylation in vitro, however not all thiol groups seem to be
 CC nitrosylated in vivo (By similarity).
 CC -!- PTM: Acetylation of Lys-545 by ARD1 increases interaction with VHL
 CC and stimulates subsequent proteasomal degradation (By similarity).
 CC -!- PTM: Regulates phosphorylation for DNA-binding (By similarity).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -!- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
 CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U59496; AAC52730.1; -;
 DR EMBL; AF003695; AAC53455.1; -;
 DR EMBL; Y09085; CAA70306.1; -;
 DR EMBL; Y13556; CAA70306.1; JOINED.
 DR EMBL; Y09085; CAA70305.1; -;
 DR EMBL; AF004155; AAC53461.1; -;
 DR EMBL; AF004141; AAC53461.1; JOINED.
 DR EMBL; AF004142; AAC53461.1; JOINED.
 DR EMBL; AF004143; AAC53461.1; JOINED.
 DR EMBL; AF004144; AAC53461.1; JOINED.
 DR EMBL; AF004145; AAC53461.1; JOINED.
 DR EMBL; AF004146; AAC53461.1; JOINED.
 DR EMBL; AF004147; AAC53461.1; JOINED.
 DR EMBL; AF004148; AAC53461.1; JOINED.
 DR EMBL; AF004149; AAC53461.1; JOINED.
 DR EMBL; AF004150; AAC53461.1; JOINED.
 DR EMBL; AF004151; AAC53461.1; JOINED.
 DR EMBL; AF004152; AAC53461.1; JOINED.
 DR EMBL; AF004153; AAC53461.1; JOINED.
 DR EMBL; AF004154; AAC53461.1; JOINED.
 DR EMBL; AK034087; BAC28578.1; -;
 DR EMBL; AK076395; BAC36320.1; -;
 DR EMBL; AK033471; BAC28305.1; -;
 DR EMBL; AK017853; BAB30975.1; -;
 DR EMBL; BC026139; AAH26139.1; -;
 DR EMBL; X95580; CAA64483.1; -;
 DR EMBL; X95002; CAA64458.1; -;
 DR PIR; JC4837; JC4837.
 DR TRANSFAC; T04666; -;
 DR MGI; MGI:106918; Hif1a.
 DR GO; GO:0009434; C:flagellum (sensu Eukarya); IDA.

Query Match 82.9%; Score 34; DB 1; Length 836;
 Best Local Similarity 87.5%; Pred. No. 3.3e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
 Db 574 MLAPVPM 581
 |||||
 |||||

RESULT 25
 Y4UG RHISN STANDARD; PRT; 71 AA.
 AC P55671;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 05-JUL-2004 (Rel. 44, Last annotation update)
 GN Hypothetical 7.8 kDa protein y4UG.
 OS ORFNames=y4UG;
 GN Rhizobium sp. (strain NGR234).
 OC Plasmid sym pNGR234a.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
 OX NCBI TaxID=394;
 RN [1]_
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
 RA Perret X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";
 RL Nature 387:394-401 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96389014; PubMed=8796346;
 RA Freiberg C., Perret X., Broughton W.J., Rosenthal A.;

```

RT "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp.
RT NGR234 using dye terminators and a thermostable 'sequenase': a
RT beginning.";
RT Genome Res. 6:590-600 (1996).
CC -1- SIMILARITY: None obvious.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68203; -; NOT ANNOTATED CDS.
DR EMBL; AE000099; AAB91879.1; -.
KW Hypothetical protein; Plasmid; Transmembrane.
FT TRANSMEM 12 32 Potential.
SQ SEQUENCE 71 AA; 7769 MW; 6552FDDA1049001 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 71;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIPTM 8
DB 21 LAPTIPL 27

RESULT 26
Q8MUT9 PRELIMINARY; PRT; 219 AA.
ID Q8MUT9
AC Q8MUT9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glutathione peroxidase.
OS Aplysia californica (California sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidia;
OC Aplysioides; Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE FROM N.A.
RA Panchin Y.V., Bodnarova M., Moroz L.L.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF510851; AAM44290.1; -.
DR HSSP; P30041; 1PRX.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR InterPro; IPR000866; AhpC-TSA.
DR Pfam; PF00578; AhpC-TSA; 1.
KW Peroxidase.
SQ SEQUENCE 219 AA; 24167 MW; 77ED48C4AB3565B5 CRC64;

Query Match 80.5%; Score 33; DB 2; Length 219;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPTM 8
DB 184 MVLPTIPTM 191

RESULT 27
Q7VNP5 PRELIMINARY; PRT; 267 AA.
ID Q7VNP5
AC Q7VNP5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=HD0928;
OS Haemophilus ducreyi.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AS017153; AAP95811.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 267 AA; 30341 MW; FA140701F47C9E77 CRC64;

Query Match 80.5%; Score 33; DB 2; Length 267;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIPTM 8
DB 81 VLAPTIPI 88

RESULT 28
Y4WD RHISN STANDARD; PRT; 377 AA.
ID Y4WD RHISN
AC P556B2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical transport protein Y4WD.
DE ORFNames=Y4WD;
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RA Freiberg C.A., Fellay R., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- FUNCTION: Could be involved in a transport system.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: TO R.MELILOTI MOSC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000103; AAB91911.1; -.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS50850; MFS; 1.
KW Hypothetical protein; Plasmid; Transmembrane; Transport.
FT TRANSMEM 4 24 Potential.
FT TRANSMEM 41 61 Potential.
FT TRANSMEM 85 105 Potential.
FT TRANSMEM 134 154 Potential.
FT TRANSMEM 159 179 Potential.
FT TRANSMEM 192 212 Potential.
FT TRANSMEM 278 298 Potential.
FT TRANSMEM 301 321 Potential.
FT TRANSMEM 327 347 Potential.
FT TRANSMEM 356 376 Potential.
SQ SEQUENCE 377 AA; 39051 MW; 49CF6E44AA0D74BD CRC64;

Query Match 80.5%; Score 33; DB 1; Length 377;

```



```

Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIPTM 8
Db 90 LAPTIPL 96
|||||:

RESULT 29
Q9ATN6 PRELIMINARY; PRT; 553 AA.
AC Q9ATN6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sesquiterpene cyclase.
GN Name=PSC2;
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
SEQUENCE FROM N.A.
RA Kim J.-B., Ha S.-H., Lee S.-K., Lee M., Lee J.-Y.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326117; AAK15641.1; -.
DR HSSP; Q40577; 5EAU.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene_synth; 1.
DR Pfam; PF03936; Terpene_synth; C; 1.
DR PIR; PF03936; Terpene_synth; C; 1.
SQ SEQUENCE 553 AA; 63946 MW; E774122F96A2394F CRC64;

Query Match 80.5%; Score 33; DB 2; Length 553;
Best Local Similarity 87.5%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPTM 8
Db 294 MLAKTIPTM 301
|||||:

RESULT 30
Q6D3T4 PRELIMINARY; PRT; 554 AA.
AC Q6D3T4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=ECA22660;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Murgall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
```

```

Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
RL EMBL; BX950851; CAG75560.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 554 AA; 63461 MW; 9879A1947A73BECF CRC64;

Query Match 80.5%; Score 33; DB 2; Length 554;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIPTM 8
Db 45 LLAPTIPL 52
|||||:

RESULT 31
Q7XQ14 PRELIMINARY; PRT; 1604 AA.
AC Q7XQ14;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OSJNB0065L13.5 protein.
GN Name=OSJNB0065L13.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Chang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL606607; CAE03362.1; -.
DR HSSP; Q91836; 1DI2.
DR Gramene; Q7XQ14; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR005034; DUF283.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR000209; Pept_s8_S53.
DR InterPro; IPR000215; Prot_inf_serpin.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00035; dsrm; 2.
DR Pfam; PF03368; DUF283; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00358; DSRW; 2.
DR SMART; SM00490; HELICc; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS50137; DS_RBD; 1.
```

Wed Feb 9 06:57:59 2005

us-10-032-361-4.rup

```

DR PROSITE: PS00821; PAZ; 1.
DR PROSITE: PS00517; RNASE_3_1; 1.
DR PROSITE: PS00142; RNASE_3_2; 2.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
KW ATP-binding; Helicase; Helicase; 181134 MW; CAE7912E4FCC9345 CRC64;
SQ SEQUENCE 1604 AA; 181134 MW;
Query Match 80.5%; Score 33; DB 2; Length 1604;
Best Local Similarity 85.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIPTM 8
Db 82 LAPTIPL 88

RESULT 32
Q82622 PRELIMINARY; PRT; 145 AA.
AC Q82622;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV7033;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=11572948; DOI=10.1073/pnas.2114333198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531 (2003).
DR EMBL: AP005048; BAC74744.1; -.
KW Complete proteome.
SQ SEQUENCE 145 AA; 15089 MW; 1415406F5247185B CRC64;

Query Match 78.0%; Score 32; DB 2; Length 145;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPTM 8
Db 6 MLAAATVPM 13

RESULT 33
Q7Y4Y5 PRELIMINARY; PRT; 191 AA.
AC Q7Y4Y5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gp53 baseplate wedge subunit.
OS Bacteriophage RB69.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;

OC T4-like viruses.
OX NCBI_TaxID=12353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98215164; PubMed=9555879;
RA Yeh L.S., Hsu T., Karam J.D.;
RT "Divergence of a DNA replication gene cluster in the T4-related
RT bacteriophage RB69.";
RL J. Bacteriol. 180:2005-2013 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20566703; PubMed=11114936; DOI=10.1128/JB.183.1.358-366.2001;
RA Tetart P., Desplats C., Kutateladze M., Monod C., Ackermann H.W.,
RA Kirsch H.M.;
RT "Phylogeny of the major head and tail genes of the wide-ranging T4-
RT type bacteriophages.";
RL J. Bacteriol. 183:358-366 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96064739; PubMed=7592876; DOI=10.1074/jbc.270.44.26558;
RA Wang C.C., Yeh L.S., Karam J.D.;
RT "Modular organization of T4 DNA polymerase. Evidence from
RT phylogenetics.";
RL J. Biol. Chem. 270:26558-26564 (1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97358535; PubMed=9215631; DOI=10.1016/S0092-8674(00)80296-2;
RA Wang J., Sattar A.K., Wang C.C., Karam J.D., Konigsberg W.H.,
RA Seitz T.A.;
RT "Crystal structure of a pol alpha family replication DNA polymerase
RT from bacteriophage RB69.";
RL Cell 89:1087-1099 (1997).
RN [5]
RP SEQUENCE FROM N.A.
RX BORJAC J., PETROV V.M., KARAM J.D.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX PETROV V., NOLAN J., CHIN D., LETAROV A., KRISCH H.M., KARAM J.D.;
RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL: AY303349; AAP76062.1; -.
SQ SEQUENCE 191 AA; 22429 MW; 794D3D2A21115716 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APTIPM 8
Db 20 APTIPM 25

RESULT 34
VG53_BPT4 STANDARD; PRT; 196 AA.
ID VG53_BPT4
AC P16011;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Baseplate structural protein Gp53 (Baseplate wedge protein 53).
GN Name=53;
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91190815; PubMed=2488704;
RA Mosig G., Lin G.W., Franklin J., Fan W.H.;
RT "Functional relationships and structural determinants of two
RT bacteriophage T4 lysozymes: a soluble (gene e) and a baseplate-
RT associated (gene 5) protein.";
RL New Biol. 1:171-179 (1989).

```

```

RN  SEQUENCE FROM N.A.
RX  MEDLINE=89296504; PubMed=2740234;
RA  Koch T., Lamm N., Rueger W.;
RT  "Sequencing, cloning and overexpression of genes of bacteriophage T4
RL  between map positions 74,325 and 77,184.";
RN  Nucleic Acids Res. 17:4392-4392(1989).
[3]
RN  SEQUENCE FROM N.A.
RX  MEDLINE=22514363; PubMed=12626685; DOI=10.1128/MMBR.67.1.86-156.2003;
RA  Miller E.S., Kutter E., Mosig G., Arita F., Kunisawa T., Ruger W.;
RT  "Bacteriophage T4 genome.";
RL  Microbiol. Mol. Biol. Rev. 67:86-156(2003).
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; X15728; CAA33748.1; -
DR  EMBL; X14845; CAA32949.1; -
DR  EMBL; AF158101; RAD42489.1; -
DR  PIR; S25239; GPBP84.
KW  Late protein; Structural protein.
SQ  SEQUENCE 196 AA; 22966 MW; F59E2E42PFE95B00 CRC64;

Query Match      78.0%; Score 32; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy  3 APTIPM 8
Db  20 APTIPM 25

RESULT 35
Q8PR29 PRELIMINARY; PRT; 243 AA.
AC  Q8PR29;
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE  Hypothetical protein XAC0137.
GN  OrderedLocusNames:XAC0137;
OS  Xanthomonas axonopodis (pv. citri).
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC  Xanthomonadaceae; Xanthomonas.
OX  NCBI_TaxID=92829;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=306 / ATCC 13902 / XV 101;
RX  MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA  da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
RA  Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA  Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA  Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA  Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA  Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA  Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA  Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA  Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA  Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA  Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA  Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA  Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA  Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA  Setubal J.C., Kitajima J.P.;
RT  "Comparison of the genomes of two Xanthomonas pathogens with differing
RT  host specificities.";
RL  Nature 417:459-463(2002).

```

```

DR  EMBL; AE011637; AAM35029.1; -
SQ  SEQUENCE 243 AA; 25706 MW; 499188BEFB31882F CRC64;

Query Match      78.0%; Score 32; DB 2; Length 243;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy  1 MLAPTIPM 8
Db  120 VLAPTUPI 127

RESULT 36
Q92LJ9 PRELIMINARY; PRT; 250 AA.
AC  Q92LJ9;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  PUTATIVE OXIDOREDUCTASE PROTEIN (EC 1.-.-.-).
GN  ORFNames=SMC02486;
OS  Rhizobium meliloti (Sinorhizobium meliloti).
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC  Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX  NCBI_TaxID=382;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=1021;
RX  MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA  Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA  Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA  Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA  Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA  Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT  Analysis of the chromosome sequence of the legume symbiont
RT  Sinorhizobium meliloti strain 1021.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC  -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC  (SDR) family.
DR  EMBL; AL591792; CAC47628.1; -
DR  HSSP; P97852; 1GZ6.
DR  GO; GO:0016491; P:oxidoreductase activity; IEA.
DR  GO; GO:0008152; P:metabolism; IEA.
DR  InterPro; IPR002198; ADH_short.
DR  InterPro; IPR002347; ADH_short C2.
DR  InterPro; IPR001395; Aldo/ket_red.
DR  Pfam; PF00106; adh_short; 1.
DR  PRINTS; PR00081; GDRHDH.
DR  PRINTS; PR00080; SDRFAMILY.
DR  ProDom; PD000288; Aldo/ket_red; 1.
DR  PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW  Complete proteome; Oxidoreductase.
SQ  SEQUENCE 250 AA; 25570 MW; 8E05010CBFB2CD5D CRC64;

Query Match      78.0%; Score 32; DB 2; Length 250;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy  2 LAPITPM 8
Db  209 LAPITPI 215

RESULT 37
Q9M3C6 PRELIMINARY; PRT; 277 AA.
AC  Q9M3C6;
DT  01-OCT-2000 (TrEMBLrel. 15, Created)
DT  01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  Hypothetical protein T26112.130 (Hypothetical protein At3g55250).
GN  Name=T26112.130; Synonyms=At3g55250;

```

```

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Monfort A., Casacuberta E., Puigdomenech P., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Quetier F., Salanoubat M.;
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132954; CAB75759.1; -
DR EMBL; AY099793; AAM20644.1; -
DR EMBL; AY128903; AAM91309.1; -
DR PIR; T47664; T47664.
KW Hypothetical protein.
SQ SEQUENCE 277 AA; 31245 MW; 7E45AB6F803AC84B CRC64;

Query Match 78.0%; Score 32; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTIPM 8
DB 180 APTIPM 185

RESULT 38
PUR3 VIGUN STANDARD; PRT; 312 AA.
AC P52423; Q9XGS3;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Phosphoribosylglycinamide formyltransferase, chloroplast precursor
DE (EC 2.1.2.2) (GART) (GAR transformylase) (5'-phosphoribosylglycinamide
DE transformylase).
GN Name=PUR3;
OS Vigna unguiculata (Cowpea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
OX NCBI_TaxID=3917;
RN [1]
RP SEQUENCE FROM N.A.
RA Hall D.J., Bussell J.D., Mann A.J., Goggins D.E., Atkins C.A.,
RA Smith P.M.C.;
RA "vupur3 mRNA from cowpea nodules encoding glycineamide ribonucleotide
RA transformylase.";
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + N(1)-(5-phospho-D-
ribosyl)glycinamide = tetrahydrofolate + N(2)-(5-formyl-N(1)-(5-
phospho-D-ribose)glycinamide.
-!- PATHWAY: De novo purine biosynthesis; third step.
-!- SUBCELLULAR LOCATION: Chloroplast.
-!- SIMILARITY: Belongs to the GART family.
-----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U10875; AAA75367.2; -
DR EMBL; AF160196; AAD45353.2; -
DR EMBL; AY189137; AAO25114.1; -
DR EMBL; AY189138; AAO25115.1; -
DR InterPro; IPR002376; formyl transf.
DR InterPro; IPR001555; GART AS.
DR InterPro; IPR004607; PurN_trans.
DR Pfam; PF00551; Formyl_trans_N; 1.
DR PROSITE; PS00373; GART; 1.
KW Chloroplast; Purine biosynthesis; Transferase; Transit peptide.
FT TRANSIT 1 73 Chloroplast (potential).
FT CHAIN 74 312 Phosphoribosylglycinamide
FT ACT_SITE 247 247 By similarity.
FT ACT_SITE 312 AA; 34582 MW; 51196B04E3D1385D CRC64;
SQ SEQUENCE 312 AA; 34582 MW; 51196B04E3D1385D CRC64;

Query Match 78.0%; Score 32; DB 1; Length 312;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIPM 8
DB 16 LAPTIPM 22

RESULT 39
Q9NKD1 PRELIMINARY; PRT; 335 AA.
AC Q9NKD1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein BG:DS05899.4;
GN Name=nacr-alpha-34E; Synonyms=BG:DS05899.4;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkeley;
RA MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Celniker S., Rubin G.M.;
RA "An exploration of the sequence of a 2.9-Mb region of the genome of
RA Drosophila melanogaster: the Adh region.";
RA Genetics 153:179-219 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkeley;
RA Celniker S.E., Agbayan A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,

```

RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
 RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003408; AAF44855.1; -;
 DR HSP: P22770; 1KL8
 DR FlyBase; FBgn0028875; nAcR-alpha-34E.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0045211; C:postsynaptic membrane; IEA.
 DR GO: GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
 DR GO: GO:0005216; F:ion channel activity; IEA.
 DR GO: GO:0030594; F:neurotransmitter receptor activity; IEA.
 DR GO: GO:0004889; F:nicotinic acetylcholine-activated cation-se. . .; IEA.
 DR GO: GO:0006811; F:ion transport; IEA.
 DR InterPro: IPR006201; Neur_chanL.
 DR InterPro: IPR006202; Neur_chanLBD.
 DR InterPro: IPR006029; Neu Channel memb.
 DR InterPro: IPR002394; Nic7ace_receptox.
 DR Pfam: PF02931; Neur_chanLBD; 1.
 DR Pfam: PF02932; Neur_chan memb; 1.
 DR PRINTS: PR00254; NICOTINICR.
 DR PRINTS: PR00252; NRIONCHANNEL.
 KW Hypothetical protein.
 SQ SEQUENCE 335 AA; 38324 MW; CC75273987E67EFC CRC64;
 Query Match 78.0%; Score 32; DB 2; Length 335;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LAPTIMP 8
 DB 53 LAPTIMP 59
 ID Q8WE47 PRELIMINARY; PRT; 343 AA.
 AC Q8WE47;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE NADH dehydrogenase subunit 2.
 GN Name=ND2;
 OS Draco blanfordii.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Draconinae;
 OC Draco.
 OX NCBI_TaxID=89021;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McGuire J.A.;
 RT "Phylogenetic systematics of Southeast Asian flying lizards (Iguania: Agamidae; Draco) as inferred from mitochondrial DNA sequences.";
 RL Biol. J. Linn. Soc. Lond. 72:203-229(2001).
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).
 CC -!- SIMILARITY: Belongs to the complex I subunit 2 family.
 DR EMBL: AF288242; AAL60090.1; -;
 DR GO: GO:0005739; C:mitochondrion; IEA.
 DR GO: GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0042773; P:ATP synthase coupled electron transport; IEA.
 DR GO: GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
 DR Pfam: PF08444; NADH dehy S2 C; 1.
 DR Pfam: PF00361; Oxidored q1; 1.
 DR PRINTS: PR01436; NADHDHGNASE2.
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 SQ SEQUENCE 343 AA; 37357 MW; 0E0E6010A1B167B5 CRC64;
 Query Match 78.0%; Score 32; DB 2; Length 343;
 Best Local Similarity 62.5%; Pred. No. 3.4e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 DB 211 MMAPTFFPL 218
 ID Q7QAJ5 PRELIMINARY; PRT; 375 AA.
 AC Q7QAJ5;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE AGCP7854 (Fragment).
 GN Name=agCG45861; ORFNames=ENSANGG00000017760;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
 OX NCBI_TaxID=180454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
 DR EMBL: AAB01008888; EAA08817.1; -;
 DR HSSP: P13297; 1IG7.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0003700; P:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR001356; Homeobox.
 DR InterPro: IPR009057; Homeodomain_like.
 DR Pfam: PF00046; Homeobox; 1.
 DR PRINTS: PR00024; HOMEBOX.
 DR ProDom: PD00010; Homeobox; 1.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PROSITE: PS00071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ NON_TER 1
 FT SEQUENCE 375 AA; 39761 MW; 314AAC3F98048430 CRC64;
 Query Match 78.0%; Score 32; DB 2; Length 375;
 Best Local Similarity 62.5%; Pred. No. 3.8e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAPTIPM 8
 DB 5 MTSFTVPM 12
 ID Q9MSF3 PRELIMINARY; PRT; 376 AA.
 AC Q9MSF3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Maturase K (Fragment).
 GN Name=matK;
 OS Nepenthes pervillei.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Nepenthaceae; Nepenthes.
 OX NCBI_TaxID=122313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Meimberg H., Ditttrich P., Bringmann G., Schlauer J., Heubl G.;
 RT "Molecular phylogeny of Caryophyllidae s.l. based on matK sequences

RT with special emphasis on carnivorous taxa. ;
RL Plant Biol. 2:218-228(2000).
DR EMBL; AF204837; AAF76391.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0008380; P:RNA splicing; IEA.
DR InterPro; IPR008998; Agglutinin.
DR InterPro; IPR000442; Intron_maturase2.
DR InterPro; IPR002866; MatK_N.
DR Pfam; PF01348; Intron_maturase2; 1.
DR Pfam; PF01824; MatK_N; 1.
KW Chloroplast.
FT NON_TER 1 1
FT NON_TER 376 376
SQ SEQUENCE 376 AA; 45033 MW; 7B2435F4BF7520E7 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 376;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAPTIPTM 8
DB 239 LVPTIPTM 245

RESULT 43
Q98H39 PRELIMINARY; PRT; 386 AA.
AC Q98H39;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE M13044 protein.
GN OrderedLocusNames=ml13044;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003001; BAB50027.1; -.
DR InterPro; IPR009334; DUF993.
DR Pfam; PF06187; DUF993; 1.
KW Complete proteome.
SQ SEQUENCE 386 AA; 41767 MW; 34CC44E534F3E334 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 386;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTPTM 8
DB 301 LMAPTPTM 308

RESULT 44
Q9RY74 PRELIMINARY; PRT; 401 AA.
AC Q9RY74;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Molybdopterin biosynthesis MoeA.
GN OrderedLocusNames=DR0076;

OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Bisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001871; AAF09669.1; -.
DR PIR; B75562; B75562.
DR HSSP; P12281; 1G8L.
DR TIGR; DR0076; -.
DR GO; GO:0006777; P:Mo-molybdopterin cofactor biosynthesis; IEA.
DR InterPro; IPR001453; MoeF_biosynth.
DR InterPro; IPR005111; MoeA_C.
DR InterPro; IPR005110; MoeA_N.
DR Pfam; PF00994; MoeF_biosynth; 1.
DR Pfam; PF03454; MoeA_C; 1.
DR Pfam; PF03453; MoeA_N; 1.
DR ProDom; PD002460; MoeF_biosynth; 1.
DR TIGRFAMS; TIGR00177; molyb_syn; 1.
KW Complete proteome.
SQ SEQUENCE 401 AA; 41504 MW; 24A3D0A56F2AD73F CRC64;

Query Match 78.0%; Score 32; DB 2; Length 401;
Best Local Similarity 75.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLAPTPTM 8
DB 1 MTAPTPTM 8

RESULT 45
Q8Y1D4 PRELIMINARY; PRT; 435 AA.
ID Q8Y1D4
AC Q8Y1D4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 25, Last annotation update)
DE PROBABLE GLUTARYL-COA DEHYDROGENASE OXIDOREDUCTASE PROTEIN (EC
DE 1.3.99.7).
GN Name=gcdH; Synonyms=RS05097; OrderedLocusNames=RS0756;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -!- SIMILARITY: Belongs to the acyl-CoA dehydrogenase family.
DR EMBL; AL646060; CAD14286.1; -.
DR HSSP; Q06319; 1BUC.
DR GO; GO:0003995; F:acyl-CoA dehydrogenase activity; IEA.

DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR Pfam: PF00441; Acyl-CoA dh; 1.
 DR Pfam: PF02770; Acyl-CoA dh.M; 1.
 DR Pfam: PF02771; Acyl-CoA dh.N; 1.
 DR PROSITE: PS00073; ACYL-CoA_DH_2; UNKNOWN 1.
 KW Complete proteome; FAD; Flavoprotein; Oxidoreductase.
 SQ SEQUENCE 435 AA; 47932 MW; 4670598BB3C0CDDD CRC64;

Query Match 78.0%; Score 32; DB 2; Length 435;
 Best Local Similarity 85.7%; Pred. No. 4.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 7
 |||||
 Db 102 MLGPTIP 108

RESULT 46

Q9XEV9 PRELIMINARY; PRT; 447 AA.
 AC Q9XEV9;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Elongation factor 1-alpha.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Mundree S.G., Singh N.P., Singh N.K.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF120093; AAD28440.1; -

DR HSP; P02994; IF60.
 DR GO: GO:0005737; C:cytoplasm; IEA.
 DR GO: GO:0005525; F:GTP binding; IEA.
 DR GO: GO:0003746; P:translation elongation factor activity; IEA.
 DR GO: GO:0006412; P:protein biosynthesis; IEA.
 DR GO: GO:0006414; P:translational elongation; IEA.
 DR InterPro; IPR004539; EFT_alpha.
 DR InterPro; IPR004160; EFTU Cterm.
 DR InterPro; IPR004161; EFTU D2.
 DR InterPro; IPR009001; Elong init C.
 DR InterPro; IPR00795; ProtSyn GTBind.
 DR InterPro; IPR009000; Translat_factor.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR Pfam; PF03143; GTP_EFTU_D3; 1.
 DR PRINTS; PR00315; ELONGATNCT.
 DR TIGRfam; TIGR00483; EF-1 alpha; 1.
 DR PROSITE; PS00301; EFACTOR-GTP; 1.
 KW Elongation factor; GTP-binding; Protein biosynthesis.
 SQ SEQUENCE 447 AA; 49065 MW; AFD1E31D7864BEDC CRC64;

Query Match 78.0%; Score 32; DB 2; Length 447;
 Best Local Similarity 62.5%; Pred. No. 4.5e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 8
 |||||
 Db 390 WMPTIP 397

RESULT 47

Q8AVC6 PRELIMINARY; PRT; 471 AA.
 AC Q8AVC6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE MG52635 protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative.";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC041717; AAH41717.1; -
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0005215; P:transporter activity; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR PROSITE; PS00850; MFS; 1.
 SQ SEQUENCE 471 AA; 51058 MW; 1504C5B99AA70B01 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 471;
 Best Local Similarity 75.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 8
 |||||
 Db 291 MLEPTIP 298

RESULT 48

Q6DJR1 PRELIMINARY; PRT; 484 AA.
 AC Q6DJR1;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Slc18a2-prov protein.
 GN Name=slc18a2-prov;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

```

DR InterPro; IPR000442; Intron_maturese2.
DR InterPro; IPR002866; MatK_N.
DR Pfam; PF01348; Intron_maturase2; 1.
DR Pfam; PF01824; MatK_N; 1.
DR Chloroplast.
SQ SEQUENCE 504 AA; 60290 MW; 30A12EF71A131945 CRC64;
Query Match 78.0%; Score 32; DB 2; Length 504;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 MLPTIPM 8
Db 364 LVPTIEM 370
| | | | |
RESULT 50
ID Q6FBU5 PRELIMINARY; PRT; 504 AA.
AC Q6FBU5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Amidase (EC 3.5.1.4).
Name=amda; OrderedLocustNames=ACIAD1618;
OS Acinetobacter sp. (Strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
NCBI_TaxID=62977;
[1]
SEQUENCE FROM N.A.
Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Ortiz S.,
Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
"Unique features revealed by the genome sequence of Acinetobacter sp.
ADP1 a versatile and naturally transformation competent bacterium.";
Nucleic Acids Res. 0:0-0(2004)).
EMBL; CR543861; CAG68466.1; -.
GO; GO:0004040; F:amidase activity; IEA.
InterPro; IPR000120; Amidase.
Pfam; PF01425; Amidase; 1.
PROSITE; PS00571; AMIDASER; 1.
Complete proteome.
KW SEQUENCE 504 AA; 54831 MW; C4DB62C86E48BA14 CRC64;
Query Match 78.0%; Score 32; DB 2; Length 504;
Best Local Similarity 62.5%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAPTIPM 8
Db 418 LMPTVDM 425
| | | | |
Search completed: February 9, 2005, 06:00:09
Job time : 26.193 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 05:55:37 ; Search time 4.63158 Seconds
(without alignments)
166.193 Million cell updates/sec

Title: US-10-032-361-4
Perfect score: 41
Sequence: 1 MLAPTIPM 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 65 summaries

Database : PIR 79: *
1: piri: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	92.7	321	2 F82073	sigma-E factor reg
2	35	85.4	462	2 G87500	hypothetical prote
3	34	82.9	810	2 JC4837	hypoxia-inducible
4	34	82.9	811	2 JC7619	hypoxia-inducible
5	34	82.9	813	2 JC3809	hypoxia-inducible
6	34	82.9	826	2 I38972	hypoxia-inducible
7	32	78.0	196	1 GPBP84	baseplate protein
8	32	78.0	277	2 T47664	hypothetical prote
9	32	78.0	312	2 T11574	phosphoribosylglyc
10	32	78.0	401	2 B75562	molybdopterin bios
11	32	78.0	666	1 A36026	kinesin-related pr
12	32	78.0	710	2 A96540	hypothetical prote
13	32	78.0	779	2 C86335	hypothetical prote
14	32	78.0	1089	2 C70522	probable mmp18 pro
15	32	78.0	3624	2 AD0835	large repetitive p
16	31	75.6	63	2 F83293	hypothetical prote
17	31	75.6	188	2 E75365	kanamycin resistan
18	31	75.6	221	1 B40335	histone H1-3 [vali
19	31	75.6	250	2 T33551	hypothetical prote
20	31	75.6	255	2 C75527	conserved hypotet
21	31	75.6	261	2 T40482	apoptosis specific
22	31	75.6	287	1 A45168	probable 3-oxoacyl
23	31	75.6	387	2 JC7398	double C2 protein,
24	31	75.6	394	2 A35745	lactate 2-monooxyg
25	31	75.6	395	2 A54949	syndecan precursor
26	31	75.6	412	2 C69188	ammonium transport
27	31	75.6	420	2 T47998	pectinacylestera
28	31	75.6	570	2 C81094	prolyl-tRNA synthe
29	31	75.6	570	2 D81847	proline-tRNA ligas

30	31	75.6	606	2 A70960	probable pckA prot
31	31	75.6	609	2 F87237	phosphoenolpyruvat
32	31	75.6	650	2 D71203	probable ATP-depen
33	31	75.6	651	2 T32875	hypothetical prote
34	31	75.6	752	2 D75219	ATP-dependent RNA
35	31	75.6	806	2 T35640	probable sensor ki
36	31	75.6	837	2 G82542	DNA ligase XF2556
37	31	75.6	1232	2 S40766	hypothetical prote
38	31	75.6	1363	2 T47492	copia-like polypor
39	30	73.2	88	2 H84352	hypothetical prote
40	30	73.2	129	2 S20611	calpastatin - rat
41	30	73.2	166	2 T17659	hypothetical prote
42	30	73.2	237	2 A11181	carboxylesterase h
43	30	73.2	237	2 AB1539	carboxylesterase h
44	30	73.2	317	1 C70356	conserved hypotet
45	30	73.2	433	2 F75566	probable benzoate
46	30	73.2	462	2 H75315	conserved hypotet
47	30	73.2	587	2 AH3472	sulfite reductase
48	30	73.2	603	2 S15074	calpastatin - rat
49	30	73.2	636	2 T03439	probable ethylene-
50	30	73.2	654	2 T10772	calpastatin - rat
51	30	73.2	782	2 H95306	probable amino aci
52	30	73.2	784	2 T20074	hypothetical prote
53	30	73.2	874	2 D75307	leucyl-tRNA synthe
54	29	70.7	53	2 H82563	hypothetical prote
55	29	70.7	124	2 C69885	phage-related prot
56	29	70.7	144	2 H83933	heat shock protein
57	29	70.7	200	1 R3MD4	ribosomal protein
58	29	70.7	201	2 S73140	ribosomal protein
59	29	70.7	201	2 A30833	ribosomal protein
60	29	70.7	201	2 T06900	ribosomal protein
61	29	70.7	202	1 R3LV4	ribosomal protein
62	29	70.7	202	2 AB2148	30S ribosomal prot
63	29	70.7	205	2 B83495	conserved hypotet
64	29	70.7	205	2 S41002	hypothetical prote
65	29	70.7	217	2 T44705	hypothetical prote

ALIGNMENTS

RESULT 1

F82073
sigma-E factor regulatory protein RseB VC2465 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82073
R:Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: F82073
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <HEI>
A:Cross-references: UNIPROT:Q9KPA8; GB:AE004316; GB:AE003852; NID:g9657034; PIDN:AAF9560
A:Experimental source: serogroup O1, strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2465
A:Map position: 1
C:Superfamily: conserved hypothetical protein HI0630

Query Match 92.7%; Score 38; DB 2; Length 321;
Best Local Similarity 87.5%; Pred. No. 3.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
DB 106 MVAPTIPM 113

RESULT 2

GB7500
 Hypothetical protein CC2028 [imported] - Caulobacter crescentus
 C/Species: Caulobacter crescentus
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C/Accession: GB7500
 R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonitskii, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A/Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A/Reference number: A87249; MUID:21173698; PMID:11259647
 A/Accession: GB7500
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-462 <STO>
 A/Cross-references: UNIPROT:Q9A6Q6; GB:AE005673; NID:gi13423503; PIDN:AAK24003.1; GSPDB:G000000000
 C/Genetics:
 A/Gene: CC2028

Query Match 85.4%; Score 35; DB 2; Length 462;

Best Local Similarity 85.7%; Pred. No. 23;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MLPTIPM 8

|||||

Db 423 MLPTVPM 429

RESULT 3

JC4837
 hypoxia-inducible factor 1 alpha - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
 C/Accession: JC4837
 R/Wenger, R.H.; Rolfs, A.; Marti, H.H.; Guenet, J.L.; Gassmann, M. Biochem. Biophys. Res. Commun. 223, 54-59, 1996
 A/Title: Nucleotide sequence, chromosomal assignment and mRNA expression of mouse hypoxia-inducible factor 1.
 A/Reference number: JC4837; MUID:96254028; PMID:8660378
 A/Accession: JC4837
 A/Molecule type: mRNA
 A/Residues: 1-810 <WEN>
 A/Cross-references: UNIPROT:O61221; EMBL:X95580; NID:gi430864; PIDN:CAA64833.1; PID:9437
 A/Comment: This factor is involved in the oxygen-regulated transcription of several genes.
 C/Genetics:
 A/Gene: Hif1alpha
 A/Map position: 12
 C/Keywords: transcription factor
 F/5-58/Region: helix-loop-helix #status predicted

Query Match 82.9%; Score 34; DB 2; Length 810;

Best Local Similarity 87.5%; Pred. No. 68;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8

|||||

Db 548 MLAPYIPM 555

RESULT 4

JC7619
 hypoxia-inducible factor 1 alpha - chicken
 C/Species: Gallus gallus (chicken)
 C/Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
 C/Accession: JC7619
 R/Takahashi, T.; Sugishita, Y.; Nojiri, T.; Shimizu, T.; Yao, A.; Kinugawa, K.; Harada, C. Biochem. Biophys. Res. Commun. 281, 1057-1062, 2001
 A/Title: Cloning of hypoxia-inducible factor 1 alpha cDNA from chick embryonic ventricular myocytes.
 A/Reference number: JC7619; MUID:21134360; PMID:11237772
 A/Contents: Embryonic ventricular myocytes
 A/Accession: JC7619
 A/Molecule type: mRNA
 A/Residues: 1-811 <TAK>

A/Cross-references: DDBJ:AB013746
 C/Comment: This factor belongs to the basic helix-loop-helix-Per/ARNT/Sim (HLH-PAS) family of transcription factors, and in redox stimuli.

C/Genetics:

A/Gene: hif-1alpha
 C/Keywords: embryo; transcription factor
 F/106-156/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS1>
 F/249-299/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS2>
 F/762-811/Domain: conserved carboxy-terminal transactivation element #status predicted <PAS3>
 F/767-768/Region: conserved dileucine repeat, important for oxygen-dependent degradation

Query Match 82.9%; Score 34; DB 2; Length 811;

Best Local Similarity 87.5%; Pred. No. 68;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8

|||||

Db 559 MLAPYIPM 566

RESULT 5

JC5809
 hypoxia-inducible factor 1 alpha - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 26-Aug-1999
 C/Accession: JC5809
 R/Ladoux, A.; Frelin, C. Biochem. Biophys. Res. Commun. 240, 552-556, 1997
 A/Title: Cardiac expressions of HIF-1 alpha and HLF/EPAS, two basic loop helix/PAS domain proteins, in the heart of the rat.
 A/Reference number: JC5809; MUID:98063274; PMID:9398602
 A/Accession: JC5809
 A/Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-813 <LAD>
 C/Comment: This protein associates to the aryl hydrocarbon receptor nuclear translocator as erythropoietin, vascular endothelial growth factor, the GLUT1 glucose transporter, and hypoxanthine phosphoribosyl transferase.

Query Match 82.9%; Score 34; DB 2; Length 813;

Best Local Similarity 87.5%; Pred. No. 68;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8

|||||

Db 548 MLAPYIPM 555

RESULT 6

I38972
 hypoxia-inducible factor 1 alpha - human
 N/Alternate names: ARNT interacting protein
 C/Species: Homo sapiens (man)
 C/Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
 C/Accession: I38972; G01875
 R/Wang, G.L.; Jiang, B.H.; Rue, E.A.; Semenza, G.L. Proc. Natl. Acad. Sci. U.S.A. 92, 5510-5514, 1995
 A/Title: Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by oxygen.
 A/Reference number: I38972; MUID:95296340; PMID:7539918
 A/Accession: I38972
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-826 <RES>
 A/Cross-references: UNIPROT:Q16665; EMBL:U22431; NID:9881345; PIDN:AA50152.1; PID:9881345
 A/Note: Parts of this sequence were confirmed by peptide sequencing
 R/Hogenesch, J.B.; Chan, W.K.; Carver, L.A.; Bradfield, C.A. submitted to the EMBL Data Library, June 1995
 A/Reference number: H00692
 A/Accession: G01875
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-826 <HOG>
 A/Cross-references: EMBL:U29165; NID:gi144012; PIDN:AA51210.1; PID:gi144013
 C/Genetics:

A:Gene: GDB:HIFIA
A:Cross-references: GDB:512229
A:Map position: 14q21-14q24
C:Keywords: heterodimer

Query Match 82.9%; Score 34; DB 2; Length 826;
Best Local Similarity 87.5%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8

Db 561 MLAPYIPM 568

RESULT 7

baseplate protein gp53 - phase T4

N:Alternate names: gene 50.1 protein; gene 53 protein
C:Species: phase T4
C:Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S25239; S04608; JF0062; S10107
R:Mosig, G.; Lin, G.W.; Franklin, J.; Fan, W.H.
New Biol. 1, 171-179, 1989
A:Title: Functional relationships and structural determinants of two bacteriophage T4 ly
A:Reference number: S25239; MUID:91190815; PMID:2488704
A:Accession: S25239
A:Molecule type: DNA
A:Residues: 1-196 <MOS>
A:Cross-references: UNIPROT:P16011; GB:X15728; NID:gl5224; PIDN:CAA33748.1; PID:gl5226
R:Koch, T.; Lamm, N.; Rueger, W.
Nucleic Acids Res. 17, 4392, 1989
A:Title: Sequencing, cloning and overexpression of genes of bacteriophage T4 between map
A:Reference number: S04608; MUID:89296504; PMID:2740234
A:Accession: S04608
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-196 <KOC>
A:Cross-references: EMBL:X14845; NID:gl5218; PIDN:CAA32949.1; PID:gl5219
C:Genetics:
A:Gene: 53; 50.1
A:Map position: 76.4-77.0
A:Superfamily: phase T4 gene 53 protein
C:Keywords: baseplate; late protein

Query Match 78.0%; Score 32; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APTIPM 8

Db 20 APTIPM 25

RESULT 8

T47664

hypothetical protein T26112.i30 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47664
R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24471
A:Accession: T47664
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-277 <MON>
A:Cross-references: UNIPROT:Q9M3C6; EMBL:AL132954
A:Experimental source: cultivar Columbia; BAC clone T26112
C:Genetics:
A:Map position: 3
A:Introns: 116/3; 171/3
A>Note: T26112.i30

Query Match 78.0%; Score 32; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APTIPM 8

Db 180 APTIPM 185

RESULT 9

T11574

phosphoribosylglycinamide formyltransferase (EC 2.1.2.2) - cowpea

C:Species: Vigna unguiculata (cowpea)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Sep-1999
C:Accession: T11574
R:Hall, D.J.; Atkins, C.A.; Smith, P.M.C.; Mann, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z17289
A:Accession: T11574
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-312 <HAL>
A:Cross-references: EMBL:U30875; NID:984817; PID:984818
A:Experimental source: strain Vita 3; root nodule; clone pCARF5
C:Genetics:
A:Gene: pur3
C:Function:

A:Description: transfers a formyl group of 10-formyltetrahydrofolate to 5'-phosphoribosyl
A:Pathway: purine nucleotide biosynthesis
C:Superfamily: phosphoribosylglycinamide formyltransferase; phosphoribosylglycinamide fo
C:Keywords: purine nucleotide biosynthesis; transferase

Query Match 78.0%; Score 32; DB 2; Length 312;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIPM 8

Db 16 LAPSIPTM 22

RESULT 10

B75562

molybdopterin biosynthesis MoeA - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2004
C:Accession: B75562
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036996; PMID:10567266
A:Accession: B75562
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-401 <WHI>
A:Cross-references: UNIPROT:Q9RY74; GB:AE001871; GB:AE000513; NID:ig6457740; PIDN:AAF0566

A:Experimental source: strain R1

C:Superfamily: Molybdenum cofactor molybdenum incorporation protein MoeA

Query Match 78.0%; Score 32; DB 2; Length 401;
Best Local Similarity 75.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8

Db 1 MTAPTIPM 8

```
RESULT 11
A36026
kinesin-related protein noda - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
C/Accession: A36026
R/Zhang, P.; Knowles, B.A.; Goldstein, L.S.B.; Hawley, R.S.
Cell 62, 1053-1062, 1990
A/Title: A kinesin-like protein required for distributive chromosome segregation in Dros
A/Reference number: A36026; MUID:90381763; PMID:2144792
A/Accession: A36026
A/Molecule type: mRNA
A/Residues: 1-666 <ZHA>
A/Cross-references: UNIPROT:P18105; GB:M36195; NID:g157779; PIDN:AAA28653.1; PID:g157780
C/Genetics:
A/Gene: FlyBase:nod
A/Cross-references: FlyBase:FBgn0002948
C/Suprafamily: kinesin-related protein noda; kinesin motor domain homology
C/Keywords: ATP; microtubule binding; mitosis; nucleotide binding; P-loop
F:9-326/Domain: kinesin motor domain homology <KWOT>
F:87-94/Region: nucleotide-binding motif A (P-loop)
F:93/Binding site: ATP (Lys) #status predicted

Query Match 78.0%; Score 32; DB 1; Length 666;
Best Local Similarity 71.4%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 0;

QY 1 MLAPTIP 7
Db 472 MVAPTP 478

RESULT 12
A96540
hypothetical protein F1413.4 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana [mouse-ear cress]
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: A96540
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: A96540
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-710 <STO>
A/Cross-references: UNIPROT:Q9SX53; GB:AE005173; NID:g5734786; PIDN:AAD50051.1; GSPDB:GN
C/Genetics:
A/Gene: F1413.4
A/Map position: 1
C/Suprafamily: prolyl oligopeptidase

Query Match 78.0%; Score 32; DB 2; Length 710;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 8
Db 592 MLDPTIP 599

RESULT 13
A86335
hypothetical protein T20H2.7 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana [mouse-ear cress]
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
```

```
C/Accession: C86335
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: C86335
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-779 <STO>
A/Cross-references: UNIPROT:Q9LNU0; GB:AE005172; NID:g8778983; PIDN:AAF9898.1; GSPDB:GN
C/Genetics:
A/Map position: 1
C/Suprafamily: subtilisin-like proteinase ag12; subtilisin homology

Query Match 78.0%; Score 32; DB 2; Length 779;
Best Local Similarity 85.7%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0;

QY 1 MLAPTIP 7
Db 489 MLAPSIP 495

RESULT 14
C70522
probable mmpL8 protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: C70522
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
C.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squires, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: C70522
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-1089 <COL>
A/Cross-references: UNIPROT:O07800; GB:Z97188; GB:AL123456; NID:g3261805; PIDN:CAB10022
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: mmpL8

Query Match 78.0%; Score 32; DB 2; Length 1089;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIP 7
Db 56 LLAPTVP 62

RESULT 15
AD0835
large repetitive protein [imported] - Salmonella enterica subsp. enterica serovar Typhi
C/Species: Salmonella enterica subsp. enterica serovar Typhi
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AD0835
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
```

A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AD0835

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-3624 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05867.1; PID:gl6503842; GSPDB:GN00176

C:Genetics:

A:Gene: STY2875

Query Match 78.0%; Score 32; DB 2; Length 3624;

Best Local Similarity 71.4%; Pred. No. 8.5e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIPTM 8

Db 2367 LAPTVPL 2373

RESULT 16

F83293

hypotheical protein PA2808 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C:Species: *Pseudomonas aeruginosa*

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C:Accession: F83293

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho

A:Reference number: AB2950; MUID:20437337; PMID:10984043

A:Accession: F83293

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-63 <STO>

A:Cross-references: UNIPROT:Q9I035; GB:AE004708; GB:AE004890; PIDN:AAG0619

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA2808

Query Match 75.6%; Score 31; DB 2; Length 63;

Best Local Similarity 71.4%; Pred. No. 18;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIP 7

Db 14 VLAPTVP 20

RESULT 17

E75365

kanamycin resistance protein-related protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: E75365

R:White, O.; Eise, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75365

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-188 <WHI>

A:Cross-references: UNIPROT:Q9RTQ4; GB:AE002011; GB:AE000513; NID:g6459461; PIDN:AAF1125

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1702

A:Map position: 1

Query Match 75.6%; Score 31; DB 2; Length 188;

Best Local Similarity 75.0%; Pred. No. 58;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIPM 8

Db 1 MLSTPTPPM 8

RESULT 18

B40335

histone H1-3 [validated] - human

N:Alternate names: histone H1.3; histone H1.c.

C:Species: *Homo sapiens* (man)

C>Date: 03-Mar-2000 #sequence_revision 03-Mar-2000 #text_change 09-Jul-2004

C:Accession: B40335; JX0086

R:Albig, W.; Kardalinos, E.; Drabant, B.; Zimmer, A.; Doenecke, D.

Genomics 10, 940-948, 1991

A:Title: Isolation and characterization of two human H1 histone genes within clusters of

A:Reference number: A40335; MUID:92009931; PMID:1918625

A:Accession: B40335

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-221 <ALB>

A:Cross-references: UNIPROT:P16402; GB:M60747; NID:gl84071; PIDN:AAA63186.1; PID:gl84072

R:One, Y.; Hayashi, H.; Iwai, K.

J. Biochem. 106, 844-857, 1989

A:Title: Human spleen histone H1; isolation and amino acid sequences of three minor vari

A:Reference number: JX0080; MUID:90130391; PMID:2613692

A:Accession: JX0086

A:Molecule type: protein

A:Residues: 2-221 <OHE>

C:Genetics:

A:Gene: GDB:H1F3; H1T

A:Cross-references: GDB:120029; OMIM:142210

A:Map position: 6p22.3-6p21.3

C:Superfamily: histone H1

C:Keywords: acetylated amino end; chromosomal protein; DNA binding; nucleosome; nucleus

F:2-221/Product: histone H1-3 #status experimental <MAT>

F:2-35/Domain: amino-terminal <ATD>

F:36-110/Domain: globular <GLB>

F:111-221/Domain: carboxyl-terminal <CTD>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental

Query Match 75.6%; Score 31; DB 1; Length 221;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIPT 7

Db 7 LAPTIPT 12

RESULT 19

T33551

hypotheical protein T1004.11 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T33551

R:Ledwith, J.; Kramer, J.; Hawkins, M.

submitted to the EMBL Data Library, October 1998

A:Description: The sequence of *C. elegans* cosmid T1004.

A:Reference number: 221366

A:Accession: T33551

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-250 <LED>

A:Cross-references: UNIPROT:Q9TZE7; EMBL:AF098997; PIDN:AAC68721.1; GSPDB:GN00020; CESP

A:Experimental source: strain Bristol N2; clone T1004

C:Genetics:

A:Gene: CESP:T1004.11

A:Map position: 2

A:introns: 40/1; 131/1; 178/3; 246/2

Query Match

75.6%; Score 31; DB 2; Length 250;

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A45168
 R:Callahan, H.L.; Beverley, S.M.
 J. Biol. Chem. 267, 24165-24168, 1992
 A:Title: A member of the aldoketo reductase family confers methotrexate resistance in Le:
 A:Reference number: A45168; MUID:93077490; PMID:1339441
 A:Accession: A45168
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-287 <GB>
 A:Cross-references: GB:I01699; NID:gl59347; PIDN:AAA29249.1; PID:gl59348
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: NAD; oxidoreductase
 F:7-224/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 75.6%; Score 31; DB 1; Length 287;
 Best Local Similarity 62.5%; Pred. No. 91;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
 |||||:
 Db 1 MRAPTVPV 8

RESULT 23
 JC7398
 double C2 protein, gamma isoform - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
 C:Accession: JC7398
 R:Fukuda, M.; Mikoshiba, K.
 Biochem. Biophys. Res. Commun. 276, 626-632, 2000
 A:Title: Doc2gamma, a third isoform of double C2 protein, lacking calcium-dependent phos
 A:Reference number: JC7398
 A:Contents: Brain
 A:Accession: JC7398
 A:Molecule type: mRNA
 A:Residues: 1-387 <FUK>
 A:Cross-references: UNIPROT:Q9ESN1; DBJ:AB046665
 C:Comment: This protein, a third isoform of double C2 protein family, functions as a neg
 C:Genetics:
 A:Gene: doc2-gamma

Query Match 75.6%; Score 31; DB 2; Length 387;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 7
 |||||:
 Db 57 MLAPAIP 63

RESULT 24
 A35745
 lactate 2-monoxygenase (EC 1.13.12.4) - Mycobacterium smegmatis
 C:Species: Mycobacterium smegmatis
 C:Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 09-Jul-2004
 C:Accession: A35745
 R:Giegel, D.A.; Williams Jr., C.H.; Massey, V.
 J. Biol. Chem. 265, 6626-6632, 1990
 A:Title: L-lactate 2-monoxygenase from Mycobacterium smegmatis. Cloning, nucleotide seq
 A:Reference number: A35745; MUID:90216683; PMID:2324094
 A:Accession: A35745
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-394 <GIE>
 A:Cross-references: UNIPROT:P21795; GB:J05402; NID:G623158; PIDN:AAA60429.1; PID:G623159
 C:Superfamily: alpha-hydroxy acid dehydrogenase, FMN-dependent; (S)-2-hydroxy-acid oxida
 C:Keywords: monoxygenase; oxidoreductase
 F:22-336/Domain: (S)-2-hydroxy-acid oxidase homology <2HY>
 F:291/Active site: His #status predicted

Query Match 75.6%; Score 31; DB 2; Length 394;

Best Local Similarity 85.7%; Pred. No. 79;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 7
 |||||:
 Db 178 MLAPTIP 184

RESULT 20
 C75527
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: C75527
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: C75527
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-255 <WHI>
 A:Cross-references: UNIPROT:Q9RXE5; GB:AE001897; GB:AE000513; NID:G6458045; PIDN:AAF0994
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR0368
 A:Map position: 1

Query Match 75.6%; Score 31; DB 2; Length 255;
 Best Local Similarity 71.4%; Pred. No. 80;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIPL 253
 |||||:
 Db 247 LAPTIPL 253

RESULT 21
 T40482
 apoptosis specific protein homolog - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T40482
 R:Beck, A.; Reinhardt, R.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, May 1997
 A:Reference number: Z21932
 A:Accession: T40482
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-261 <BEC>
 A:Cross-references: UNIPROT:O74971; EMBL:AL023706; PIDN:CAAL19290.1; GSPDB:GN00067; SPDB:
 A:Experimental source: strain 972h; cosmid c4B4
 C:Genetics:
 A:Gene: SPDB:SPBC4B4.10c
 A:Map position: 2 73/3; 73/2; 131/3
 A:Introns: 33/3; 73/2; 131/3

Query Match 75.6%; Score 31; DB 2; Length 261;
 Best Local Similarity 62.5%; Pred. No. 82;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
 |||||:
 Db 52 LLAPSIPL 59

RESULT 22
 A45168
 probable 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) - Leishmania major
 N:Alternate names: H region methotrexate resistance protein
 C:Species: Leishmania major

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A45168
 R:Callahan, H.L.; Beverley, S.M.
 J. Biol. Chem. 267, 24165-24168, 1992
 A:Title: A member of the aldoketo reductase family confers methotrexate resistance in Le:
 A:Reference number: A45168; MUID:93077490; PMID:1339441
 A:Accession: A45168
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-287 <GB>
 A:Cross-references: GB:I01699; NID:gl59347; PIDN:AAA29249.1; PID:gl59348
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: NAD; oxidoreductase
 F:7-224/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 75.6%; Score 31; DB 1; Length 287;
 Best Local Similarity 62.5%; Pred. No. 91;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
 |||||:
 Db 1 MRAPTVPV 8

RESULT 23
 JC7398
 double C2 protein, gamma isoform - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
 C:Accession: JC7398
 R:Fukuda, M.; Mikoshiba, K.
 Biochem. Biophys. Res. Commun. 276, 626-632, 2000
 A:Title: Doc2gamma, a third isoform of double C2 protein, lacking calcium-dependent phos
 A:Reference number: JC7398
 A:Contents: Brain
 A:Accession: JC7398
 A:Molecule type: mRNA
 A:Residues: 1-387 <FUK>
 A:Cross-references: UNIPROT:Q9ESN1; DBJ:AB046665
 C:Comment: This protein, a third isoform of double C2 protein family, functions as a neg
 C:Genetics:
 A:Gene: doc2-gamma

Query Match 75.6%; Score 31; DB 2; Length 387;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 7
 |||||:
 Db 57 MLAPAIP 63

RESULT 24
 A35745
 lactate 2-monoxygenase (EC 1.13.12.4) - Mycobacterium smegmatis
 C:Species: Mycobacterium smegmatis
 C:Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 09-Jul-2004
 C:Accession: A35745
 R:Giegel, D.A.; Williams Jr., C.H.; Massey, V.
 J. Biol. Chem. 265, 6626-6632, 1990
 A:Title: L-lactate 2-monoxygenase from Mycobacterium smegmatis. Cloning, nucleotide seq
 A:Reference number: A35745; MUID:90216683; PMID:2324094
 A:Accession: A35745
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-394 <GIE>
 A:Cross-references: UNIPROT:P21795; GB:J05402; NID:G623158; PIDN:AAA60429.1; PID:G623159
 C:Superfamily: alpha-hydroxy acid dehydrogenase, FMN-dependent; (S)-2-hydroxy-acid oxida
 C:Keywords: monoxygenase; oxidoreductase
 F:22-336/Domain: (S)-2-hydroxy-acid oxidase homology <2HY>
 F:291/Active site: His #status predicted

Query Match 75.6%; Score 31; DB 2; Length 394;

```
Best Local Similarity 71.4%; Pred. No. 1.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 2;

QY 2 LAPTIM 8
Db 19 VAPLPM 25

RESULT 25
A:Accession: A54949
A:Title: syndecan precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
A:Accession: A54949
R:Spring, J.; Paine-Saunders, S.E.; Hynes, R.O.; Bernfield, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 3334-3338, 1994
A:Title: Drosophila syndecan: conservation of a cell-surface heparan sulfate proteoglycan
A:Reference number: A54949; MUID:94211858; PMID:8159748
A:Accession: A54949
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-395 <SPR>
A:Cross-references: UNIPROT:P49415; GB:U03282; NID:g437282; PIDN:AAC34307.1; PID:g437283
A:Note: Parts of this sequence were confirmed by protein sequencing
C:Genetics:
A:Gene: FlyBase:Scd
A:Cross-references: FlyBase:FBgn0010415
C:Keywords: transmembrane protein

Query Match 75.6%; Score 31; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIM 7
Db 165 LAPTIM 170

RESULT 26
A:Accession: C69188
A:Title: ammonium transporter - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
A:Accession: C69188
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcu
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: C69188
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-412 <MTH>
A:Cross-references: UNIPROT:Q26759; GB:AE000846; GB:AE000666; NID:g2621740; PIDN:AAB8516
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH663
C:Superfamily: ammonium transporter nrgA

Query Match 75.6%; Score 31; DB 2; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIM 7
Db 100 LAPTIM 105

RESULT 27
A:Accession: T47998
A:Title: pectinacetyltransferase-like protein T17J13.20 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
```

```
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
A:Accession: T47998
R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer,
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24482
A:Accession: T47998
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-420 <RIE>
A:Cross-references: UNIPROT:Q9M1R8; EMBL:AL1138651
A:Experimental source: cultivar Columbia; BAC clone T17J13
C:Genetics:
A:Map position: 3
A:Introns: 69/1; 96/3; 140/1; 164/3; 202/1; 236/2; 256/3; 274/3; 300/3; 336/1; 384/3
A:Note: T17J13.20

Query Match 75.6%; Score 31; DB 2; Length 420;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAPTIM 8
Db 47 LMPSVPM 54

RESULT 28
A:Accession: C81094
A:Title: prollyl-tRNA synthetase NMB1339 [imported] - Neisseria meningitidis (strain MC58 serogro
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
A:Accession: C81094
R:Tetelin, H.; Saunders, N.J.; Heideberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.P.;
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: C81094
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-570 <TEN>
A:Cross-references: UNIPROT:Q9J214; GB:AE002482; GB:AE002098; NID:g7226577; PIDN:AAF4171
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1339
C:Superfamily: proline-tRNA ligase

Query Match 75.6%; Score 31; DB 2; Length 570;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIM 8
Db 235 LAPLPL 241

RESULT 29
A:Accession: D81847
A:Title: proline-tRNA ligase (EC 6.1.1.15) NMA1553 [imported] - Neisseria meningitidis (strain Z2
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
A:Accession: D81847
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D81847
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-570 <PAR>
A:Cross-references: UNIPROT:Q9JU09; GB:ALI62756; GB:ALI57959; NID:g7380091; PIDN:CAB8478
```

Wed Feb 9 06:57:58 2005

us-10-032-361-4.rpr

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: p068; NMA1553

C;Superfamily: proline-trna ligase

C;Keywords: ligase

Query Match 75.6%; Score 31; DB 2; Length 570;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAPTIMP 8

Db 235 LAPTLPL 241

RESULT 30

A70960

probable pckA protein - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: A70960

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: A70960

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-606 <COL>

A;Cross-references: UNIPROT:P96393; GB:Z92669; GB:AL123456; NID:G3242271; PIDN:CAB07006.

A;Experimental source: strain H37RV

C;Genetics:

A;Gene: pckA

C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 75.6%; Score 31; DB 2; Length 606;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 7

Db 280 MLQPTIP 286

RESULT 31

F87237

phosphoenolpyruvate carboxykinase [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: F87237

R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho

; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S

A;Title: Massive gene decay in the leprosy bacillus.

A;Reference number: A86909; MUID:21128732; PMID:11234002

A;Accession: F87237

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-609 <STO>

A;Cross-references: UNIPROT:O06084; GB:AL450380; NID:gl3093848; PIDN:CAC32156.1; GSPDB:G

C;Genetics:

A;Gene: pckA

C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 75.6%; Score 31; DB 2; Length 609;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 7

Db 280 MLQPTIP 286

RESULT 32

D71203

probable ATP-dependent RNA helicase - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004

C;Accession: D71203

R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: D71203

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-650 <KAW>

A;Cross-references: UNIPROT:O59524; GB:AP000007; NID:G3236134; PIDN:EAA31019.1; PID:G325

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH1896

C;Keywords: ATP; nucleotide binding; P-loop

F;33-40/Region: nucleotide-binding motif A (P-loop)

F;131-136/Region: nucleotide-binding motif B

F;135-138/Region: DEAD/H motif

Query Match 75.6%; Score 31; DB 2; Length 650;

Best Local Similarity 75.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 8

Db 60 MLAPTIP 67

RESULT 33

T32875

hypothetical protein W05G11.6 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T32875

R;Davidson, S.; Langston, Y.; O'Neal, D.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of C. elegans cosmid W05G11.

A;Reference number: Z21238

A;Accession: T32875

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-651 <DAV>

A;Cross-references: UNIPROT:O44906; EMBL:AF040660; PIDN:AAC71149.1; GSPDB:GN00021; CESP

A;Experimental source: strain Bristol N2; clone W05G11

C;Genetics:

A;Gene: CESP:W05G11.6

A;Map position: 3

A;Introns: 92/2; 222/1; 284/3; 579/3

C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 75.6%; Score 31; DB 2; Length 651;

Best Local Similarity 71.4%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIP 7

Db 313 MLPTVP 319

RESULT 34

D75219

ATP-dependent RNA helicase, eif-4a family PAB0190 - Pyrococcus abyssi (strain Orsay)

C:Species: *Pyrococcus abyssi*
 C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C:Accession: D75219
 R:anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure
 A:Reference number: A75001
 A:Accession: D75219
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-752 <KAW>
 A:Cross-references: UNIPROT:Q9V1Z5; GB:AJ248283; GB:AL096836; NID:G5457433; PIDN:CAB4920
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB0190
 C:Superfamily: ATP-dependent RNA helicase eIF-4A
 F:33-40/Region: nucleotide-binding motif A (P-loop) #status atypical
 F:131-136/Region: nucleotide-binding motif B
 F:135-138/Region: DEAD/H motif

Query Match 75.6%; Score 31; DB 2; Length 752;
 Best Local Similarity 71.4%; Pred. No. 2.5e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLPTIPM 8
 Db 60 MLPTKPL 67

RESULT 35
 T35640
 C:Species: *Streptomyces coelicolor*
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C:Accession: T35640
 R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A:Reference number: Z21584
 A:Accession: T35640
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-806 <SER>
 A:Cross-references: UNIPROT:Q8CJR8; EMBL:AL079356; PIDN:CAB45635.1; GSPDB:GN000070; SCOE
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC6G9.44c

Query Match 75.6%; Score 31; DB 2; Length 806;
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MLPTIPM 8
 Db 523 LAPVFPV 529

RESULT 36
 G82542
 DNA ligase XF2556 [imported] - *Xylella fastidiosa* (strain 9a5c)
 C:Species: *Xylella fastidiosa*
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 12-Jul-2004
 C:Accession: G82542
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A:Reference number: A82515; PMID:10910347
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: G82542
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-837 <SIM>
 A:Cross-references: UNIPROT:Q9PAG2; GB:AE004063; GB:AE003849; NID:G9107766; PIDN:AAF8535
 A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, P.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.P.P.; Camargo, A.A.; Camargo, L.E.A.; Carrato, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
 J.D.; Junqueira, M.L.; Kemper, S.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, P.
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaas
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; V
 A:Reference number: A59328
 C:Contents: annotation
 C:Genetics:
 A:Gene: XP2556
 C:Superfamily: NAD+-dependent DNA ligase, Liga type

Query Match 75.6%; Score 31; DB 2; Length 837;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIMP 8
 Db 694 LAPTLPL 700

RESULT 37
 S40766
 hypothetical protein ZKS12.5 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
 C:Accession: S40766
 R:Hawkins, T.; Ainscough, R.
 submitted to the EMBL Data Library, February 1993
 A:Reference number: S40759
 A:Accession: S40766
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1232 <HAW>
 A:Cross-references: UNIPROT:P34643; EMBL:Z22177; NID:G297989; PID:G297994
 C:Genetics:
 A:Introns: 5/3; 19/1; 52/2; 108/2; 306/3; 344/2; 490/3; 542/3; 634/3; 703/2; 1060/3; 111

Query Match 75.6%; Score 31; DB 2; Length 1232;
 Best Local Similarity 62.5%; Pred. No. 4.3e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLPTIPM 8
 Db 1188 MMPPTMPM 1195

RESULT 38
 T47492
 copia-like polypeptide - *Arabidopsis thaliana*
 N:Alternate names: protein F9K21.100
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
 C:Accession: T47492
 R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.
 submitted to the Protein Sequence Database, February 2000
 A:Reference number: Z24467
 A:Accession: T47492
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1363 <JOR>
 A:Cross-references: UNIPROT:Q9M1F5; EMBL:AL138657
 A:Experimental source: cultivar Columbia; BAC clone F9K21
 C:Genetics:
 A:Map position: 3
 A>Note: F9K21.100
 C:Superfamily: retrovirus-related polypeptide

Wed Feb 9 06:57:58 2005

Query Match 75.6%; Score 31; DB 2; Length 1363;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 3 APTIPM 8
Db 514 APTVPM 519

RESULT 39
H84352
Hypothetical protein Vng2024h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84352
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabcig, Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: H84352
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-88 <STO>
A;Cross-references: UNIPROT:Q9HNN2; GB:AE004437; NID:g10581453; PIDN:AAG20188.1; GSPDB:G
C;Genetics:
A;Gene: VNG2024H

Query Match 73.2%; Score 30; DB 2; Length 88;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LAPTIP 7
Db 10 LAPTVP 15

RESULT 40
S20611
calpastatin - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S20611
R;Lee, W.J.; Hatanaka, M.; Maki, M.
Biochim. Biophys. Acta 1129, 251-253, 1992
A;Title: Multiple forms of rat calpastatin cDNA in the coding region of functionally unkn
A;Reference number: S20610; MUID:92110395; PMID:1730065
A;Accession: S20611
A;Molecule type: mRNA
A;Residues: 1-129 <LEE>
A;Cross-references: UNIPROT:P27321; EMBL:X62520; NID:g57530; PIDN:CAA44386.1; PID:g57531
C;Superfamily: calpain inhibitor

Query Match 73.2%; Score 30; DB 2; Length 129;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LAPTIPM 8
Db 113 LTFTLPM 119

RESULT 41
T17659
Hypothetical protein A168R - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17659
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999

A;Reference number: Z18806
A;Accession: T17659
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-166 <GRA>
A;Cross-references: UNIPROT:Q84488; EMBL:U42580; NID:g4028896; PIDN:AAC96536.1
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Note: A168R
C;Superfamily: Chlorella virus PBCV-1 hypothetical protein A168R

Query Match 73.2%; Score 30; DB 2; Length 166;
Best Local Similarity 71.4%; Pred. No. 81;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MLAPTIP 7
Db 126 MITPTIP 132

RESULT 42
A11181
carboxylesterase homolog lmo0857 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: A11181
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: A11181
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <GLA>
A;Cross-references: UNIPROT:Q8Y8N9; GB:NC_003210; PIDN:CAC98935.1; PID:g16410260; GSPDB
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0857
C;Superfamily: esterase/lipase

Query Match 73.2%; Score 30; DB 2; Length 237;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLAPTIP 7
Db 32 VLAPTLP 38

RESULT 43
AB1539
carboxylesterase homolog lin0850 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AB1539
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <GLA>
A;Cross-references: UNIPROT:Q92DG3; GB:AL592022; PIDN:CAC96082.1; PID:g16413301; GSPDB
A;Experimental source: strain Clip11262

C:Genetics:
A:Gene: lin0850
C:Superfamily: esterase/lipase

Query Match 73.2%; Score 30; DB 2; Length 237;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAPTIP 7
:||||:
Db 32 VLAPTLP 38

RESULT 44

C70356

conserved hypothetical protein aq_632 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: C70356

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: C70356

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-317 <AQF>

A:Cross-references: UNIPROT:O66879; GB:AE000659; GB:AE000657; NID:G2983238; PIDN:AAC0684

A:Experimental source: strain VFS

C:Genetics:

A:Gene: aq_632

C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0486

Query Match 73.2%; Score 30; DB 1; Length 317;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAPTIP 8
:||||:
Db 59 MLSPVPL 66

RESULT 45

F75566

probable benzoate membrane transport protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: F75566

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: F75566

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-433 <WHI>

A:Cross-references: UNIPROT:Q9RY94; GB:AE001868; GB:AE000513; NID:G6457707; PIDN:AAF0964

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0056

A:Map position: 1

Query Match 73.2%; Score 30; DB 2; Length 433;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIP 7
:||||:
Db 229 LAPTIP 234

RESULT 46

H75315

conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: H75315

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: H75315

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-462 <WHI>

A:Cross-references: UNIPROT:Q9RSM6; GB:AE002045; GB:AE000513; NID:G6459886; PIDN:AAF1164

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2098

A:Map position: 1

C:Superfamily: lincomycin-resistance protein lmrB

Query Match 73.2%; Score 30; DB 2; Length 462;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAPTIP 8
:||||:
Db 110 LAPTML 116

RESULT 47

AH3472

sulfite reductase (ferredoxin) (EC 1.8.7.1) [imported] - Brucella melitensis (strain 16)

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Mar-2004

C:Accession: AH3472

R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

, Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AH3472

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-587 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL52947.1; PID:G17983797; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI1766

A:Map position: 1

C:Superfamily: ferredoxin-nitrite reductase

C:Keywords: oxidoreductase

Query Match 73.2%; Score 30; DB 2; Length 587;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LAPTIP 8
:||||:
Db 558 LTPTLPM 564

RESULT 48

SI5074

calpastatin - rat

C:Species: Rattus sp. (rat)

C>Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C:Accession: SI5074

R:ishida, S.; Emori, Y.; Suzuki, K.

Biochim. Biophys. Acta 1088, 436-438, 1991

Wed Feb 9 06:57:58 2005

Job time : 7.63158 secs

A;Title: Rat calpastatin has diverged primary sequence from other mammalian calpastatins
A;Reference number: S15074; MUID:91198152; PMID:2015306
A;Accession: S15074
A;Molecule type: mRNA
A;Residues: 1-603 <ISH>
A;Cross-references: UNIPROT:P27321; EMBL:X56729; NID:G57783; PIDN:CAA40053.1; PID:G57784
C;Superfamily: calpain inhibitor

Query Match 73.2%; Score 30; DB 2; Length 603;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAPTIM 8
| | | | |
Db 62 LTPTLPM 68

RESULT 49
T03439
probable ethylene-response protein - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03439
R;Yau, C.P.; Yip, W.K.
submitted to the EMBL Data Library, July 1997
A;Reference number: Z14946
A;Accession: T03439
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-636 <YAU>
A;Cross-references: UNIPROT:O24176; EMBL:AF013979; NID:G2281704; PID:G2281705
C;Genetics:
A;Gene: ERS

Query Match 73.2%; Score 30; DB 2; Length 636;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAPTIPM 8
| | | | |
Db 444 MLAPDPL 451

RESULT 50
T10772
calpastatin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10772
R;de Tullio, R.; Sparatore, B.; Salamino, F.; Melloni, E.; Pontremoli, S.
FEBS Lett. 422, 113-117, 1998
A;Title: Rat brain contains multiple mRNAs for calpastatin.
A;Reference number: Z17133; MUID:98133328; PMID:9475181
A;Accession: T10772
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-654 <DET>
A;Cross-references: UNIPROT:O55152; EMBL:Y13588; NID:G2765339; PIDN:CAA73916.1; PID:G276
A;Experimental source: strain Sprague Dawley, brain
C;Genetics:
A;Gene: CAST
C;Superfamily: calpain inhibitor

Query Match 73.2%; Score 30; DB 2; Length 654;
Best Local Similarity 71.4%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LAPTIM 8
| | | | |
Db 113 LTPTLPM 119

Search completed: February 9, 2005, 05:57:31

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 20:10:07 ; Search time 11.6491 Seconds
(without alignments)
51.265 Million cell updates/sec

Title: US-10-032-361-4
Perfect score: 35
Sequence: 1 MLAXTTPM 8

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 65 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	85.7	650	4	US-09-583-110-5074
2	30	85.7	654	4	US-09-107-433-2950
3	29	82.9	408	4	US-09-902-540-13623
4	29	82.9	436	4	US-09-198-452A-451
5	29	82.9	732	4	US-09-438-185A-436
6	29	82.9	814	4	US-09-328-352-4373
7	28	80.0	19	4	US-09-959-873B-8
8	28	80.0	34	4	US-09-959-873B-9
9	28	80.0	56	4	US-09-438-833-13
10	28	80.0	116	4	US-09-438-833-8
11	28	80.0	175	4	US-09-328-352-6893
12	28	80.0	176	4	US-09-902-540-16036
13	28	80.0	177	4	US-09-328-352-6077
14	28	80.0	207	4	US-09-543-681A-6664
15	28	80.0	266	4	US-09-252-991A-23473
16	28	80.0	287	4	US-09-489-039A-8965
17	28	80.0	288	4	US-09-438-833-9
18	28	80.0	301	4	US-09-438-833-10
19	28	80.0	311	4	US-09-438-833-7
20	28	80.0	532	4	US-09-949-016-7389
21	28	80.0	613	4	US-09-438-833-6
22	28	80.0	648	4	US-09-328-352-7882
23	28	80.0	652	4	US-09-438-833-5
24	28	80.0	697	4	US-09-603-208A-226
25	28	80.0	756	4	US-09-438-833-11
26	28	80.0	765	4	US-09-270-767-32645
27	28	80.0	765	4	US-09-270-767-47862

28	80.0	805	2	US-08-480-473B-4	Sequence 4, Appli
28	80.0	805	3	US-08-915-213-4	Sequence 4, Appli
28	80.0	805	4	US-09-235-217-4	Sequence 4, Appli
28	80.0	805	5	PCT-US96-10251-4	Sequence 7, Appli
28	80.0	810	1	US-08-785-241-7	Sequence 12, Appli
28	80.0	813	4	US-09-438-833-12	Sequence 6, Appli
28	80.0	826	1	US-08-785-241-6	Sequence 2, Appli
28	80.0	826	2	US-08-480-473B-2	Sequence 2, Appli
28	80.0	826	3	US-08-915-213-2	Sequence 2, Appli
28	80.0	826	3	US-09-148-547-2	Sequence 2, Appli
28	80.0	826	3	US-09-235-217-2	Sequence 2, Appli
28	80.0	826	3	US-09-380-662-23	Sequence 2, Appli
28	80.0	826	4	US-09-438-833-1	Sequence 1, Appli
28	80.0	826	4	US-09-702-705-330	Sequence 330, App
28	80.0	826	4	US-09-736-457-330	Sequence 330, App
28	80.0	826	4	US-09-383-581-2	Sequence 2, Appli
28	80.0	826	4	US-09-614-124B-330	Sequence 330, App
28	80.0	826	4	US-09-671-325-330	Sequence 330, App
28	80.0	826	4	US-09-589-184-330	Sequence 330, App
28	80.0	826	4	US-09-658-824-330	Sequence 330, App
28	80.0	826	4	US-09-959-873B-18	Sequence 18, Appl
28	80.0	826	4	US-09-949-016-6089	Sequence 6089, Ap
28	80.0	826	4	US-09-967-388-4	Sequence 4, Appli
28	80.0	826	5	PCT-US96-10251-2	Sequence 2, Appli
28	80.0	827	4	US-09-919-039-149	Sequence 149, App
27	77.1	148	2	US-08-612-858-26	Sequence 26, Appl
27	77.1	148	4	US-09-107-433-3293	Sequence 3293, Ap
27	77.1	150	4	US-09-583-110-3588	Sequence 3588, Ap
27	77.1	165	4	US-09-519-232-48	Sequence 48, Appl
27	77.1	206	4	US-09-071-035-272	Sequence 272, App
27	77.1	207	4	US-09-071-035-270	Sequence 270, App
27	77.1	227	4	US-09-248-796A-19726	Sequence 19726, A
27	77.1	239	4	US-09-252-991A-19704	Sequence 19704, A
27	77.1	274	4	US-09-489-039A-9632	Sequence 9632, Ap
27	77.1	350	2	US-08-468-790-2	Sequence 2, Appli
27	77.1	379	4	US-09-252-991A-21117	Sequence 21117, A
27	77.1	389	4	US-09-543-681A-4461	Sequence 4461, Ap
27	77.1	394	3	US-08-791-115B-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1

US-09-583-110-5074
; Sequence 5074, Application US/09583110
; Patent No. 6695703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5074
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae.
US-09-583-110-5074

Query Match 85.7%; Score 30; DB 4; Length 650;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTTPM 8
|||: |||

Db 458 MLAVNIPM 465

RESULT 2
US-09-107-433-2950
; Sequence 2950, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS

NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Aziniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2950:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...654
SEQUENCE DESCRIPTION: SEQ ID NO: 2950:
US-09-107-433-2950

Query Match 85.7%; Score 30; DB 4; Length 654;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
||:|
Db 462 MLAVNIPM 469

RESULT 3
US-09-902-540-13623
; Sequence 13623, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13623
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13623

Query Match 82.9%; Score 29; DB 4; Length 408;
Best Local Similarity 71.4%; Pred. No. 3.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
||:|
Db 331 LAVTVP 337

RESULT 4
US-09-198-452A-451
; Sequence 451, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 451
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 1...436
; OTHER INFORMATION: Xaa-unknown or other
US-09-198-452A-451

Query Match 82.9%; Score 29; DB 4; Length 436;
Best Local Similarity 71.4%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
||:|
Db 148 LAVTVP 154

RESULT 5
US-09-438-185A-436
; Sequence 436, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 436
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPn0434
US-09-438-185A-436

Query Match 82.9%; Score 29; DB 4; Length 732;
Best Local Similarity 71.4%; Pred. No. 6.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MLAXTIPM 8
Db 444 LAATVPM 450

RESULT 6
US-09-328-352-4373
; Sequence 4373, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4373
; LENGTH: 814
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4373

Query Match 82.9%; Score 29; DB 4; Length 814;
Best Local Similarity 71.4%; Pred. No. 7.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIP 7
Db 746 MLATVPM 752

RESULT 7
US-09-959-873B-8
; Sequence 8, Application US/09959873B
; Patent No. 6787326
; GENERAL INFORMATION:
; APPLICANT: Ratcliffe, Peter John
; APPLICANT: Maxwell, Patrick Henry
; APPLICANT: Pugh, Christopher William
; TITLE OF INVENTION: Interaction Between the VHL Tumour
; TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods
; FILE REFERENCE: 3547.1000-000
; CURRENT APPLICATION NUMBER: US/09/959,873B
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: PCT/GB00/01826
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: GB9911047.0
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Motif
US-09-959-873B-8

Query Match 80.0%; Score 28; DB 4; Length 19;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 6 MLAPYIPM 13

RESULT 8
US-09-959-873B-9
; Sequence 9, Application US/09959873B
; Patent No. 6787326
; GENERAL INFORMATION:
; APPLICANT: Ratcliffe, Peter John
; APPLICANT: Maxwell, Patrick Henry
; APPLICANT: Pugh, Christopher William
; TITLE OF INVENTION: Interaction Between the VHL Tumour
; TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods
; FILE REFERENCE: 3547.1000-000
; CURRENT APPLICATION NUMBER: US/09/959,873B
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: PCT/GB00/01826
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: GB9911047.0
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-959-873B-9

Query Match 80.0%; Score 28; DB 4; Length 34;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 13 MLAPYIPM 20

RESULT 9
US-09-438-833-13
; Sequence 13, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 531-584 of human HIF-1 alpha
US-09-438-833-13

Query Match 80.0%; Score 28; DB 4; Length 54;
Best Local Similarity 75.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 13 MLAPYIPM 20

Db 31 MLAPYIPM 38

RESULT 10

US-09-438-833-8
; Sequence 8, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 526-641 of human HIF-1 alpha
US-09-438-833-8

Query Match 80.0%; Score 28; DB 4; Length 116;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|:|:|:
Db 36 MLAPYIPM 43

RESULT 11

US-09-328-352-6893
; Sequence 6893, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6893
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6893

Query Match 80.0%; Score 28; DB 4; Length 154;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
|:|:|:
Db 35 MAHTIPM 41

RESULT 12

US-09-902-540-16036
; Sequence 16036, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10

Query Match 80.0%; Score 28; DB 4; Length 154;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|:|:|:
Db 137 MLCNTIPL 144

; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16036
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16036

Query Match 80.0%; Score 28; DB 4; Length 176;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|:|:|:
Db 1 MVAATVPL 8

RESULT 13

US-09-328-352-6077
; Sequence 6077, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6077
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6077

Query Match 80.0%; Score 28; DB 4; Length 177;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIP 7
|:|:|:
Db 82 MLATILP 88

RESULT 14

US-09-543-681A-6664
; Sequence 6664, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709,1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6664
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6664

Query Match 80.0%; Score 28; DB 4; Length 207;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|:|:|:
Db 137 MLCNTIPL 144


```
RESULT 15
US-09-252-991A-23473
; Sequence 23473, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23473
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23473

Query Match      80.0%; Score 28; DB 4; Length 266;
Best Local Similarity 62.5%; Pred. No. 3.3e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAETIPM 8
      |||:||||
Db      163 MLAPSPM 170

RESULT 16
US-09-489-039A-8965
; Sequence 8965, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8965
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8965

Query Match      80.0%; Score 28; DB 4; Length 287;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAETIPM 8
      |||:||||
Db      4 MLAKSIPL 11

RESULT 17
US-09-438-833-9
; Sequence 9, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
US-09-438-833-9

Query Match      80.0%; Score 28; DB 4; Length 288;
Best Local Similarity 75.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAETIPM 8
      |||:||||
Db      36 MLAPYIPM 43

RESULT 18
US-09-438-833-10
; Sequence 10, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
US-09-438-833-10

Query Match      80.0%; Score 28; DB 4; Length 301;
Best Local Similarity 75.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAETIPM 8
      |||:||||
Db      36 MLAPYIPM 43

RESULT 19
US-09-438-833-7
; Sequence 7, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
US-09-438-833-7

Query Match      80.0%; Score 28; DB 4; Length 311;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7882
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7882

Query Match 80.0%; Score 28; DB 4; Length 648;
Best Local Similarity 62.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 82 MLAQDVPM 89

RESULT 23

US-09-438-833-5
; Sequence 5, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 1-652 of human HIF-1 alpha
US-09-438-833-5

Query Match 80.0%; Score 28; DB 4; Length 652;
Best Local Similarity 75.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 561 MLAQDVPM 568

RESULT 24

US-09-603-208A-226
; Sequence 226, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zeider, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; TITLE OF INVENTION: RESISTANCE AND TOLERANCE PROTEINS
; FILE REFERENCE: BGI-124CP
; CURRENT APPLICATION NUMBER: US/09/603,208A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27

QY 1 MLAXTIPM 8
DB 231 MLAQDVPM 238

US-09-949-016-7389
; Sequence 7389, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7389
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7389

Query Match 80.0%; Score 28; DB 4; Length 532;
Best Local Similarity 75.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 267 MLAQDVPM 274

RESULT 21

US-09-438-833-6
; Sequence 6, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-438-833-6

Query Match 80.0%; Score 28; DB 4; Length 613;
Best Local Similarity 75.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 348 MLAQDVPM 355

RESULT 22

US-09-328-352-7882
; Sequence 7882, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932914.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941382.7
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 226
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-603-208A-226

Query Match 80.0%; Score 28; DB 4; Length 697;
Best Local Similarity 50.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 384 LLATVPL 391

RESULT 25
US-09-438-833-11
; Sequence 11, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 71-826 of human HIF-1 alpha
US-09-438-833-11

Query Match 80.0%; Score 28; DB 4; Length 756;
Best Local Similarity 75.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 491 MLAPYIPM 498

RESULT 26
US-09-270-767-32645
; Sequence 32645, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32645
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32645

Query Match 80.0%; Score 28; DB 4; Length 765;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 151 LLVITIPM 158

RESULT 27
US-09-270-767-47862
; Sequence 47862, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47862
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47862

Query Match 80.0%; Score 28; DB 4; Length 765;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 151 LLVITIPM 158

RESULT 28
US-08-480-473B-4
; Sequence 4, Application US/08480473B
; Patent No. 5882914
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,473B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-473B-4
Query Match 80.0%; Score 28; DB 2; Length 805;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 540 MLAPYIPM 547

RESULT 29
US-08-915-213-4
; Sequence 4, Application US/08915213
; Patent No. 6020462
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,213
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-915-213-4
Query Match 80.0%; Score 28; DB 3; Length 805;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 540 MLAPYIPM 547

RESULT 30
US-09-235-217-4
; Sequence 4, Application US/09235217
; Patent No. 6222018
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/235,217
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-235-217-4
Query Match 80.0%; Score 28; DB 3; Length 805;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 540 MLAPYIPM 547

RESULT 31
PCT-US96-10251-4
; Sequence 4, Application PC/TUS9610251
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10251
```

;; FILING DATE: 06-JUN-1996
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haile, Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: 07265/053WO1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619/678-5070
;; TELEFAX: 619/678-5099
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 805 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
PCT-US96-10251-4

Query Match 80.0%; Score 28; DB 5; Length 805;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 540 MLAPYIPM 547

RESULT 32
US-08-785-241-7
; Sequence 7, Application US/08785241
; Patent No. 5695963
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; APPLICANT: Tian, Hui
; TITLE OF INVENTION: Endothelial PAS Domain Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,241
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1229
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 810 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-785-241-7

Query Match 80.0%; Score 28; DB 1; Length 810;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 548 MLAPYIPM 555

RESULT 33
US-09-438-833-12
; Sequence 12, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 1-813 of human HIF-1 alpha
US-09-438-833-12

Query Match 80.0%; Score 28; DB 4; Length 813;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 561 MLAPYIPM 568

RESULT 34
US-08-785-241-6
; Sequence 6, Application US/08785241
; Patent No. 5695963
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; APPLICANT: Tian, Hui
; TITLE OF INVENTION: Endothelial PAS Domain Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,241
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1229
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-785-241-6

Query Match      80.0%; Score 28; DB 1; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 561 MLAPYIPM 568

RESULT 35
US-08-480-473B-2
; Sequence 2, Application US/08480473B
; Patent No. 5882914
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,473B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-915-213-2

Query Match      80.0%; Score 28; DB 3; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 561 MLAPYIPM 568

RESULT 37
US-09-148-547-2
; Sequence 2, Application US/09148547
; Patent No. 6124131
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: Hypoxia Inducible Factor-1 and Methods of Use
; FILE REFERENCE: 07265/151001
; CURRENT APPLICATION NUMBER: US/09/148,547
; CURRENT FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-148-547-2

Query Match      80.0%; Score 28; DB 3; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 561 MLAPYIPM 568

RESULT 38
US-09-235-217-2
; Sequence 2, Application US/09235217
; Patent No. 6222018
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
```

; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/235,217
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/480,473
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haile, Lisa A.
 ; REGISTRATION NUMBER: 38,347
 ; REFERENCE/DOCKET NUMBER: 07265/053001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619/678-5070
 ; TELEFAX: 619/678-5099
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 826 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-235-217-2

Query Match 80.0%; Score 28; DB 3; Length 826;
 Best Local Similarity 75.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
 Db 561 MLAPYIPM 568

RESULT 39
 ; Sequence 23, Application US/09380662
 ; Patent No. 6376199
 ; GENERAL INFORMATION:
 ; APPLICANT: Caniggia, Isabella
 ; APPLICANT: Post, Martin
 ; APPLICANT: Lye, Stephen
 ; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF TROPHOBLAST
 ; FILE REFERENCE: 11757.38USWO
 ; CURRENT APPLICATION NUMBER: US/09/380,662
 ; CURRENT FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: PCT/CA98/00180
 ; PRIOR FILING DATE: 1998-03-05
 ; PRIOR APPLICATION NUMBER: US 60/039,919
 ; PRIOR FILING DATE: 1997-03-07
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 23
 ; LENGTH: 826
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-380-662-23

Query Match 80.0%; Score 28; DB 3; Length 826;
 Best Local Similarity 75.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8

Db 561 MLAPYIPM 568
 RESULT 40
 ; Sequence 1, Application US/09438833
 ; Patent No. 6436654
 ; GENERAL INFORMATION:
 ; APPLICANT: Pharmacia & Upjohn
 ; TITLE OF INVENTION: Protein variants
 ; FILE REFERENCE: 1848
 ; CURRENT APPLICATION NUMBER: US/09/438,833
 ; CURRENT FILING DATE: 1999-11-12
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 826
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
 ; VOLUME: 92
 ; PAGES: 5510-5514
 ; DATABASE ACCESSION NUMBER: GenBank U22431
 ; DATABASE ENTRY DATE: 1995-06-28
 ; US-09-438-833-1

Query Match 80.0%; Score 28; DB 4; Length 826;
 Best Local Similarity 75.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
 Db 561 MLAPYIPM 568

RESULT 41
 ; Sequence 330, Application US/09702705
 ; Patent No. 6504010
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedrick, Tom
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.478C14
 ; CURRENT APPLICATION NUMBER: US/09/702,705
 ; CURRENT FILING DATE: 2000-10-30
 ; NUMBER OF SEQ ID NOS: 1833
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 330
 ; LENGTH: 826
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-702-705-330

Query Match 80.0%; Score 28; DB 4; Length 826;
 Best Local Similarity 75.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
 Db 561 MLAPYIPM 568

```

RESULT 42
US-09-736-457-330
; Sequence 330, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-330

Query Match      80.0%; Score 28; DB 4; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAXTIPM 8
Db      561 MLAPYIPM 568

RESULT 45
US-09-671-325-330
; Sequence 330, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-325-330

Query Match      80.0%; Score 28; DB 4; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAXTIPM 8
Db      561 MLAPYIPM 568

RESULT 46
US-09-589-184-330
; Sequence 330, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.

```


; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-589-184-330

Query Match 80.0%; Score 28; DB 4; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 561 MLAPYIPM 568

RESULT 47
US-09-658-824-330
; Sequence 330, Application US/09658824
; Patent No. 6746846
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C11
; CURRENT APPLICATION NUMBER: US/09/658,824
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 1788
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-658-824-330

Query Match 80.0%; Score 28; DB 4; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 561 MLAPYIPM 568

RESULT 48
US-09-959-873B-18
; Sequence 18, Application US/09959873B
; Patent No. 6787326
; GENERAL INFORMATION:
; APPLICANT: Ratcliffe, Peter John
; APPLICANT: Maxwell, Patrick Henry
; APPLICANT: Pugh, Christopher William
; TITLE OF INVENTION: Interaction Between the VHL Tumour
; TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods

; TITLE OF INVENTION: Relating Thereto
; FILE REFERENCE: 3547.1000-000
; CURRENT APPLICATION NUMBER: US/09/959,873B
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: PCT/GB00/01826
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: GB9911047.0
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-959-873B-18

Query Match 80.0%; Score 28; DB 4; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 561 MLAPYIPM 568

RESULT 49
US-09-949-016-6089
; Sequence 6089, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6089
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6089

Query Match 80.0%; Score 28; DB 4; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 561 MLAPYIPM 568

RESULT 50
US-09-967-388-4
; Sequence 4, Application US/09967388
; Patent No. 6818430
; GENERAL INFORMATION:
; APPLICANT: JEFFEREY M. ARBEIT
; TITLE OF INVENTION: USE OF HIF-1ALPHA VARIANTS TO ACCELERATE
; TITLE OF INVENTION: WOUND HEALING
; FILE REFERENCE: UC077.001A
; CURRENT APPLICATION NUMBER: US/09/967,388
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4

Wed Feb 9 06:11:21 2005

```
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HUMAN
US-09-967-388-4

Query Match      80.0%; Score 28; DB 4; Length 826;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAXTIPM 8
      |||: |||
Db      561 MLAPYIPM 568
```

Search completed: February 8, 2005, 20:34:01
Job time : 20.6491 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 20:15:22 ; Search time 31.4386 Seconds
(without alignments)
82.880 Million cell updates/sec

Title: US-10-032-361-4

Perfect score: 35

Sequence: 1 MLAXTIPM 8

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 65 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	97.1	145	14	US-10-156-761-14561
2	32	91.4	385	16	US-10-437-963-158574
3	31	88.6	387	15	US-10-369-493-13354
4	31	88.6	517	15	US-10-108-260A-4826
5	31	88.6	654	15	US-10-297-022-21
6	30	85.7	8	14	US-10-101-662A-20
7	30	85.7	8	14	US-10-287-670-20
8	30	85.7	93	15	US-10-424-599-250147
9	30	85.7	99	16	US-10-767-701-41659
10	30	85.7	139	15	US-10-424-599-241025
11	30	85.7	525	15	US-10-424-599-282869
12	30	85.7	650	9	US-09-815-242-13341
13	30	85.7	650	10	US-09-769-787-95

14	30	85.7	650	15	US-10-282-122A-73910	Sequence 73910, A
15	30	85.7	650	17	US-10-472-928-1694	Sequence 1694, Ap
16	29	82.9	105	16	US-10-437-963-189333	Sequence 189333, A
17	29	82.9	120	14	US-10-080-170-469	Sequence 469, App
18	29	82.9	120	16	US-10-468-356-469	Sequence 469, App
19	29	82.9	120	16	US-10-369-493-12254	Sequence 12254, A
20	29	82.9	341	15	US-10-424-599-268103	Sequence 268103, A
21	29	82.9	367	15	US-10-369-493-13035	Sequence 13035, A
22	29	82.9	378	15	US-10-282-122A-48355	Sequence 48355, A
23	29	82.9	436	15	US-10-289-762-451	Sequence 451, App
24	29	82.9	528	16	US-10-437-963-116010	Sequence 116010, A
25	29	82.9	740	16	US-10-437-963-14585	Sequence 14585, A
26	29	82.9	751	14	US-10-132-134-16	Sequence 16, Appl
27	29	82.9	798	15	US-10-282-122A-45056	Sequence 45056, A
28	29	82.9	1353	16	US-10-467-685-13	Sequence 13, Appl
29	29	82.9	1646	16	US-10-437-963-183777	Sequence 183777, A
30	29	82.9	1646	16	US-10-437-963-183777	Sequence 183777, A
31	28	80.0	8	14	US-10-101-662A-16	Sequence 16, Appl
32	28	80.0	8	14	US-10-101-662A-19	Sequence 19, Appl
33	28	80.0	8	14	US-10-101-616-8	Sequence 8, Appl
34	28	80.0	8	14	US-10-287-670-16	Sequence 16, Appl
35	28	80.0	8	14	US-10-287-670-19	Sequence 19, Appl
36	28	80.0	19	16	US-10-901-583-8	Sequence 8, Appl
37	28	80.0	20	14	US-10-101-662A-15	Sequence 15, Appl
38	28	80.0	20	14	US-10-287-670-15	Sequence 15, Appl
39	28	80.0	20	14	US-10-287-670-25	Sequence 25, Appl
40	28	80.0	34	16	US-10-901-583-9	Sequence 9, Appl
41	28	80.0	54	9	US-09-922-958-5	Sequence 5, Appl
42	28	80.0	59	10	US-09-776-724A-97	Sequence 97, Appl
43	28	80.0	65	9	US-09-796-692-1600	Sequence 1600, Ap
44	28	80.0	65	9	US-09-796-692-1961	Sequence 1961, Ap
45	28	80.0	65	9	US-09-796-692-2484	Sequence 2484, Ap
46	28	80.0	65	14	US-10-040-862-1600	Sequence 1600, Ap
47	28	80.0	65	14	US-10-040-862-1961	Sequence 1961, Ap
48	28	80.0	65	14	US-10-040-862-2484	Sequence 2484, Ap
49	28	80.0	65	15	US-10-057-475B-1600	Sequence 1600, Ap
50	28	80.0	65	15	US-10-057-475B-1961	Sequence 1961, Ap
51	28	80.0	65	15	US-10-057-475B-2484	Sequence 2484, Ap
52	28	80.0	65	15	US-10-154-884B-1600	Sequence 1600, Ap
53	28	80.0	65	15	US-10-154-884B-1961	Sequence 1961, Ap
54	28	80.0	65	15	US-10-154-884B-2484	Sequence 2484, Ap
55	28	80.0	65	16	US-10-764-324-1600	Sequence 1600, Ap
56	28	80.0	65	16	US-10-764-324-1961	Sequence 1961, Ap
57	28	80.0	65	16	US-10-764-324-2484	Sequence 2484, Ap
58	28	80.0	72	15	US-10-424-599-275258	Sequence 275258, A
59	28	80.0	73	16	US-10-767-701-53697	Sequence 53697, A
60	28	80.0	82	16	US-10-437-963-145064	Sequence 145064, A
61	28	80.0	97	15	US-10-630-590-60	Sequence 60, Appl
62	28	80.0	101	15	US-10-424-599-252502	Sequence 252502, A
63	28	80.0	119	14	US-10-080-170-142	Sequence 142, App
64	28	80.0	119	16	US-10-080-170-142	Sequence 142, App
65	28	80.0	119	16	US-10-468-356-142	Sequence 142, App

ALIGNMENTS

RESULT 1
US-10-156-761-14561
; Sequence 14561, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10156.761
; CURRENT FILING DATE: 2002-05-29

us-10-032-361-4.rapb

Wed Feb 9 06:11:21 2005

```

; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14561
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14561

Query Match          97.1%; Score 34; DB 14; Length 145;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 6 MLATVPM 13

RESULT 2
US-10-437-963-158574
; Sequence 158574, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 158574
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(385)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58035C.1.pep
US-10-437-963-158574

Query Match          91.4%; Score 32; DB 16; Length 385;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 283 MIATVPM 290

RESULT 3
US-10-369-493-13354
; Sequence 13354, Application US/10369493
; Publication No. US20030233678A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B

```

```

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13354
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(387)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13354

Query Match          88.6%; Score 31; DB 15; Length 387;
Best Local Similarity 75.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 97 MLTSTIPM 104

RESULT 4
US-10-108-260A-4826
; Sequence 4826, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4826
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4826

Query Match          88.6%; Score 31; DB 15; Length 517;
Best Local Similarity 75.0%; Pred. No. 4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 245 ILATIPM 252

RESULT 5
US-10-297-022-21
; Sequence 21, Application US/10297022
; Publication No. US20030216310A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: THORNTON, Michael
; APPLICANT: WALIA, Narinder K.
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: TANG, Y. Tom
; APPLICANT: AZIMZAI, Valda
; APPLICANT: BRUNS, Christopher M.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YANG, Junming

```

```
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: SANJANWALA, Madhu S.
; APPLICANT: RAUMANN, Brigitte E.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAPALIA, April
; APPLICANT: GREENE, Barrie D.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: KEARNEY, Liam
; APPLICANT: ELLIOTT, Vicky S.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: BURFORD, Neil
; APPLICANT: DING, Li
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: HILLMAN, Jennifer L.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0109 PCT
; CURRENT APPLICATION NUMBER: US/10/297,022
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 60/208,424; 60/209,001; 60/210,588; 60/212,335; 60/213,747;
; PRIOR FILING DATE: 2000-05-26; 2000-06-01; 2000-06-08; 2000-06-16; 2000-06-22; 2000-06-22
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030216310A1 7474202CD1
US-10-027-022-21

Query Match      88.6%; Score 31; DB 15; Length 654;
Best Local Similarity 75.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 382 ILATTIPM 389

RESULT 6
US-10-101-662A-20
; Sequence 20, Application US/10101662A
; Publication No. US20030022198A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Livingston, David A
; APPLICANT: Kim, William
; TITLE OF INVENTION: Light Emitting Fusion Proteins and Diagnostic and
; TITLE OF INVENTION: Therapeutic Methods Thereof
; FILE REFERENCE: 20363-009
; CURRENT APPLICATION NUMBER: US/10/101,662A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: peptide
US-10-287-670-20

Query Match      85.7%; Score 30; DB 14; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 1 MLAPAIM 8

US-10-287-670-20
```

```
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: peptide
US-10-101-662A-20

Query Match      85.7%; Score 30; DB 14; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 1 MLAPAIM 8

RESULT 7
US-10-287-670-20
; Sequence 20, Application US/10287670
; Publication No. US20030150005A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., et al.
; TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and
; TITLE OF INVENTION: Diagnostic and
; TITLE OF INVENTION: Therapeutic Methods Thereof
; FILE REFERENCE: 20363-009C1PI
; CURRENT APPLICATION NUMBER: US/10/287,670
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 10/101,662
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,812
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,816
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: peptide
US-10-287-670-20

Query Match      85.7%; Score 30; DB 14; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 1 MLAPAIM 8

US-10-287-670-20
```


; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13341
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13341

Query Match 85.7%; Score 30; DB 9; Length 650;
Best Local Similarity 75.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 458 MLAVNIPM 465

RESULT 13
US-09-769-787-95
; Sequence 95, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129W0
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 95
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-95

Query Match 85.7%; Score 30; DB 10; Length 650;
Best Local Similarity 75.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 458 MLAVNIPM 465

RESULT 14
US-10-282-122A-73910
; Sequence 73910, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 73910
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73910

Query Match 85.7%; Score 30; DB 15; Length 650;
Best Local Similarity 75.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 458 MLAVNIPM 465

RESULT 15
US-10-472-928-1694
; Sequence 1694, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926W0
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1694
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: PTS system, fructose specific IIABC components
; OTHER INFORMATION: Cellular location: membrane
; OTHER INFORMATION: Similar to strain R6 sequence 15902824 (O.E+01)
US-10-472-928-1694

Query Match 85.7%; Score 30; DB 17; Length 650;
Best Local Similarity 75.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 458 MLAVNIPM 465

RESULT 16
US-10-437-963-189333
; Sequence 189333, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 189333
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85852C.1.pep
US-10-437-963-189333

Query Match 82.9%; Score 29; DB 16; Length 105;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIP 7
Db 53 MLASTVP 59

RESULT 17
US-10-080-170-469
; Sequence 469, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 469
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-469

Query Match 82.9%; Score 29; DB 14; Length 120;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8

Db 94 LLAGTIPL 101

RESULT 18
US-10-080-170-469
; Sequence 469, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 469
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-469

Query Match 82.9%; Score 29; DB 16; Length 120;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 94 LLAGTIPL 101

RESULT 19
US-10-468-356-469
; Sequence 469, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 469
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-469

Query Match 82.9%; Score 29; DB 16; Length 120;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 94 LLAGTIPL 101

RESULT 20
US-10-369-493-12254
; Sequence 12254, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12254
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12254

Query Match 82.9%; Score 29; DB 15; Length 248;
Best Local Similarity 71.4%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
|||:|
Db 204 LALTVPM 210

RESULT 21
US-10-424-599-268103
; Sequence 268103, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 268103
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(341)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MFT3847_84120C.1.pep
US-10-424-599-268103

Query Match 82.9%; Score 29; DB 15; Length 341;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
|||:|
Db 276 LAYTVP 282

RESULT 22
US-10-369-493-13035
; Sequence 13035, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13035
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(367)
; OTHER INFORMATION: unsure at all xaa locations
US-10-369-493-13035

Query Match 82.9%; Score 29; DB 15; Length 367;
Best Local Similarity 71.4%; Pred. No. 7.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIP 7
|||:|
Db 293 MLACTVP 299

RESULT 23
US-10-282-122A-48355
; Sequence 48355, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48355

```

; LENGTH: 378
; TYPE: PRT
; ORGANISM: Bacteroides fragilis
US-10-282-122A-48355

Query Match      82.9%; Score 29; DB 15; Length 378;
Best Local Similarity 62.5%; Pred. No. 7.7e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
   |||:||||
Db 247 LLSITIPM 254

RESULT 24
US-10-289-762-451
; Sequence 451, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 451
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 1...436
; OTHER INFORMATION: Xaa=unknown or other
US-10-289-762-451

Query Match      82.9%; Score 29; DB 15; Length 436;
Best Local Similarity 71.4%; Pred. No. 9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
   |||:||
Db 148 LARTVPM 154

RESULT 25
US-10-437-963-116010
; Sequence 116010, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116010
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_19552C.1.pap
US-10-437-963-116010

Query Match      82.9%; Score 29; DB 16; Length 528;
Best Local Similarity 75.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
   |||:||||
Db 414 MLAPPIPM 421

RESULT 26
US-10-437-963-145585
; Sequence 145585, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 145585
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46291C.1.pap
US-10-437-963-145585

Query Match      82.9%; Score 29; DB 16; Length 740;
Best Local Similarity 62.5%; Pred. No. 1.6e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
   |||:||||
Db 693 VIATIPM 700

RESULT 27
US-10-132-134-16
; Sequence 16, Application US/10132134
; Publication No. US20030171562A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Yang, Xianshu
; APPLICANT: Staiff, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES
; FILE REFERENCE: 3012-2US
; CURRENT APPLICATION NUMBER: US/10/132,134
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Streptomyces platensis subsp. rosaceus
US-10-132-134-16

Query Match      82.9%; Score 29; DB 14; Length 751;
Best Local Similarity 71.4%; Pred. No. 1.7e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
   |||:||
Db 470 LARTVPM 476
```

```
RESULT 28
US-10-282-122A-45056
; Sequence 45056, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45056
; LENGTH: 798
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-45056

Query Match      82.9%; Score 29; DB 15; Length 798;
Best Local Similarity 71.4%; Pred. No. 1.8e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy      1 MLAXTIP 7
      |||:|
Db      730 MLATVP 736

RESULT 29
US-10-467-685-13
; Sequence 13, Application US/10467685
; Publication No. US20040116666A1
; GENERAL INFORMATION:
; APPLICANT: LEE, Ernestine A.; DING, Li;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: BRUNS, Christopher M.; ELLIOTT, Vicki S.;
; APPLICANT: CHAWLA, Narinder K.; FORSYTHE, Ian J.;
; APPLICANT: RAUMANN, Brigitte E.; BURFORD, Neil;
; APPLICANT: LAL, Preeti G.; THORNTON, Michael B.;
; APPLICANT: GANDHI, Ameena R.; ARVIZU, Chandra S.;
```

```
; APPLICANT: YAO, Monique G.; YUE, Henry;
; APPLICANT: XU, Yuming; HAFALIA, April J.A.;
; APPLICANT: ISON, Craig H.; CHEN, Hui-Mei
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0356 USN
; CURRENT APPLICATION NUMBER: US/10/467,685
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/US02/03657
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,892
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/271,168
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/272,890
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/276,860
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 60/278,255
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/280,538
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/351,359
; PRIOR FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 1353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 4921451CD1
US-10-467-685-13

Query Match      82.9%; Score 29; DB 16; Length 1353;
Best Local Similarity 62.5%; Pred. No. 3.2e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy      1 MLAXTIP 8
      |||:|
Db      455 LLSVTIPM 462

RESULT 30
US-10-437-963-183777
; Sequence 183777, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 183777
; LENGTH: 1646
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80835C.1.pep
US-10-437-963-183777

Query Match      82.9%; Score 29; DB 16; Length 1646;
Best Local Similarity 62.5%; Pred. No. 4e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MLAXTIPM 8
Db      320 MIAATIPV 327

RESULT 31
US-10-101-662A-16
; Sequence 16, Application US/10101662A
; Publication No. US20030022198A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Livingston, David A
; APPLICANT: Kim, William
; TITLE OF INVENTION: Light Emitting Fusion Proteins and Diagnostic and
; FILE REFERENCE: 20363-009
; CURRENT APPLICATION NUMBER: US/10101,662A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: peptide
US-10-101-662A-16

Query Match      80.0%; Score 28; DB 14; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAXTIPM 8
Db      1 MLAPYIPM 8

RESULT 32
US-10-101-662A-19
; Sequence 19, Application US/10101662A
; Publication No. US20030022198A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Livingston, David A
; APPLICANT: Kim, William
; TITLE OF INVENTION: Light Emitting Fusion Proteins and Diagnostic and
; FILE REFERENCE: 20363-009
; CURRENT APPLICATION NUMBER: US/10101,662A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
; OTHER INFORMATION: peptide
US-10-101-662A-16

Query Match      80.0%; Score 28; DB 14; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAXTIPM 8
Db      1 MLAPYIPM 8

RESULT 33
US-10-101-816-8
; Sequence 8, Application US/10101816
; Publication No. US20030045686A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Ivan, Mircea
; TITLE OF INVENTION: Mutains of Hypoxia Inducible Factor Alpha and Methods
; FILE REFERENCE: 20363-008
; CURRENT APPLICATION NUMBER: US/10101,816
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Conserved HIF
```

OTHER INFORMATION: 8-mer
US-10-101-816-8

Query Match 80.0%; Score 28; DB 14; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 1 MLAPVIPM 8

RESULT 34

US-10-287-670-16
Sequence 16, Application US/10287670
Publication No. US20030150005A1

GENERAL INFORMATION:
APPLICANT: Kaelin Jr., et al.

TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and
TITLE OF INVENTION: Diagnostic and

TITLE OF INVENTION: Therapeutic Methods Thereof

FILE REFERENCE: 20363-009CIP1

CURRENT APPLICATION NUMBER: US/10/287,670

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 10/101,662

PRIOR FILING DATE: 2002-03-19

PRIOR APPLICATION NUMBER: 10/101,812

PRIOR FILING DATE: 2002-03-19

PRIOR APPLICATION NUMBER: 10/101,816

PRIOR FILING DATE: 2002-03-19

PRIOR APPLICATION NUMBER: 60/277,425

PRIOR FILING DATE: 2001-03-20

PRIOR APPLICATION NUMBER: 60/277,431

PRIOR FILING DATE: 2001-03-20

PRIOR APPLICATION NUMBER: 60/277,440

PRIOR FILING DATE: 2001-03-20

PRIOR APPLICATION NUMBER: 60/332,493

PRIOR FILING DATE: 2001-11-09

PRIOR APPLICATION NUMBER: 60/345,131

PRIOR FILING DATE: 2001-12-20

PRIOR APPLICATION NUMBER: 60/342,598

PRIOR FILING DATE: 2001-12-20

PRIOR APPLICATION NUMBER: 60/345,132

PRIOR FILING DATE: 2001-12-20

Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 16

LENGTH: 8

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:synthetic

OTHER INFORMATION: peptide

US-10-287-670-16

Query Match 80.0%; Score 28; DB 14; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 1 MLAPVIPM 8

RESULT 35

US-10-287-670-19

Sequence 19, Application US/10287670

Publication No. US20030150005A1

GENERAL INFORMATION:

APPLICANT: Kaelin Jr., et al.

TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and

TITLE OF INVENTION: Diagnostic and

US-10-901-593-8

Query Match 80.0%; Score 28; DB 16; Length 19;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
|||: |||
Db 6 MLAXYIPM 13

RESULT 37

US-10-101-662A-15
; Sequence 15, Application US/10101662A
; Publication No. US2003002198A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Livingston, David A
; APPLICANT: Kim, William
; TITLE OF INVENTION: Light Emitting Fusion Proteins and Diagnostic and
; FILE REFERENCE: 20363-009
; CURRENT APPLICATION NUMBER: US/10/101,662A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 15

LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (9)
OTHER INFORMATION: Wherein Xaa is hydroxyproline
US-10-101-662A-15
Query Match 80.0%; Score 28; DB 14; Length 20;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8

|||: |||
Db 6 MLAXYIPM 13

RESULT 38

US-10-287-670-15
; Sequence 15, Application US/10287670
; Publication No. US20030150005A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., et al.
; TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and
; FILE REFERENCE: 20363-009CIP1
; CURRENT APPLICATION NUMBER: US/10/287,670
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 10/101,662
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,812
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,816
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; FILE REFERENCE: 20363-009CIP1

Qy 1 MLAXTIPM 8

|||: |||
Db 6 MLAXYIPM 13

RESULT 39

US-10-287-670-25
; Sequence 25, Application US/10287670
; Publication No. US20030150005A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., et al.
; TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and
; FILE REFERENCE: 20363-009CIP1
; CURRENT APPLICATION NUMBER: US/10/287,670
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 10/101,662
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,812
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,816
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; FILE REFERENCE: 20363-009CIP1

Qy 1 MLAXTIPM 8

|||: |||
Db 6 MLAXYIPM 13

US-10-287-670-25
; Sequence 25, Application US/10287670
; Publication No. US20030150005A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., et al.
; TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and
; FILE REFERENCE: 20363-009CIP1
; CURRENT APPLICATION NUMBER: US/10/287,670
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 10/101,662
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,812
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,816
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 15

LENGTH: 20

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: VARIANT

LOCATION: (9)

OTHER INFORMATION: Wherein Xaa is hydroxyproline

US-10-287-670-15

Query Match 80.0%; Score 28; DB 14; Length 20;

Best Local Similarity 87.5%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8

|||: |||
Db 6 MLAXYIPM 13

```
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-670-25

Query Match      80.0%; Score 28; DB 14; Length 20;
Best Local Similarity 75.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MMAXTIPM 8
Db 6 MLAPYIPM 13

RESULT 40
US-10-901-583-9
; Sequence 9, Application US/10901583
; Publication No. US20050003452A1
; GENERAL INFORMATION:
; APPLICANT: Ratcliffe, Peter John
; APPLICANT: Maxwell, Patrick Henry
; APPLICANT: Pugh, Christopher William
; TITLE OF INVENTION: Interaction Between the VHL Tumour
; TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods
; TITLE OF INVENTION: Relating Thereto
; FILE REFERENCE: 3547.1000-000
; CURRENT APPLICATION NUMBER: US/10/901,583
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: US/09/959,873
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: PCT/GB00/01826
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: GB9911047.0
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-10-901-583-9

Query Match      80.0%; Score 28; DB 16; Length 34;
Best Local Similarity 75.0%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MMAXTIPM 8
Db 13 MLAPYIPM 20

RESULT 41
US-09-922-958-5
; Sequence 5, Application US/09922958
; Patent No. US20020048794A1
; GENERAL INFORMATION:
; APPLICANT: FOELLINGER, Lorenz
; APPLICANT: PEREIRA, Teresa
; APPLICANT: RUAS, Jorge
; TITLE OF INVENTION: MECHANISM OF CONDITIONAL REGULATION OF THE HYPOXIA-INDUCIBLE FACT
; TITLE OF INVENTION: THE VON HIPPEL-LINDAU TUMOR SUPPRESSOR PROTEIN
; FILE REFERENCE: 3743/49008
; CURRENT APPLICATION NUMBER: US/09/922,958
; CURRENT FILING DATE: 2001-08-07
```

```
; PRIOR APPLICATION NUMBER: US 60/223,480
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-958-5

Query Match      80.0%; Score 28; DB 9; Length 54;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MMAXTIPM 8
Db 30 MLAPYIPM 37

RESULT 42
US-09-776-724A-97
; Sequence 97, Application US/09776724A
; Publication No. US20030050455A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 64 Human Secreted Proteins
; FILE REFERENCE: P2011
; CURRENT APPLICATION NUMBER: US/09/776,724A
; CURRENT FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/180,909
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 09/669,688
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: 09/229,982
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: PCT/US98/14613
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/052,661
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,872
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,871
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,874
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,873
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,870
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/052,875
; PRIOR FILING DATE: 1997-07-16
; PRIOR APPLICATION NUMBER: 60/053,440
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,441
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,442
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/056,359
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,725
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,985
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,952
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,989
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,361
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,726
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,724
; PRIOR FILING DATE: 1997-08-18
```

; PRIOR APPLICATION NUMBER: 60/055,946
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,683
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 97
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-776-724A-97

Query Match 80.0%; Score 28; DB 10; Length 59;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
||:||||:
Db 1 MLCQTIPL 8

RESULT 43
US-09-796-692-1600
; Sequence 1600, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1600
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1600

Query Match 80.0%; Score 28; DB 9; Length 65;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
||:||||:
Db 7 MLCQTIPL 14

RESULT 44
US-09-796-692-1961
; Sequence 1961, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1961
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1961

Query Match 80.0%; Score 28; DB 9; Length 65;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
||:||||:
Db 7 MLCQTIPL 14

RESULT 45
US-09-796-692-2484
; Sequence 2484, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2484
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-2484

Query Match 80.0%; Score 28; DB 9; Length 65;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MLCXTIPM 8
|||:
Db 7 MLCQTIPL 14

RESULT 46
US-10-040-862-1600
; Sequence 1600, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416

; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1600
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-040-862-1600

Query Match 80.0%; Score 28; DB 14; Length 65;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MLCXTIPM 8
|||:
Db 7 MLCQTIPL 14

RESULT 47
US-10-040-862-1961
; Sequence 1961, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1961
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-040-862-1961

Query Match 80.0%; Score 28; DB 14; Length 65;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
||:||||
Db 7 MLCQTIPL 14

RESULT 48

US-10-040-862-2484
; Sequence 2484, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2484
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-040-862-2484

Query Match 80.0%; Score 28; DB 14; Length 65;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
||:||||
Db 7 MLCQTIPL 14

RESULT 49

US-10-057-475B-1600
; Sequence 1600, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane

APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Aijun
APPLICANT: Ordonez, Nadia
APPLICANT: Carter, Lauren
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1600
LENGTH: 65
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-475B-1600

Query Match 80.0%; Score 28; DB 15; Length 65;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
||:||||
Db 7 MLCQTIPL 14

RESULT 50

US-10-057-475B-1961
; Sequence 1961, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545

US-10-057-475B-1961

```
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1961
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-475B-1961
```

```
Query Match      80.0%; Score 28; DB 15; Length 65;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MLXTPM 8
   ||:|:|:
Db 7 MLCQTPL 14
```

```
Search completed: February 8, 2005, 20:37:59
Job time : 42.4386 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 19:40:51 ; Search time 42.807 Seconds
(without alignments)
72.280 Million cell updates/sec

Title: US-10-032-361-4
Perfect score: 35
Sequence: 1 MLAXTIPM 8

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 65 summaries

Database : A_Geneseq_16Dec04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	8	ABR82377	AbR82377 Hypoxia-i
2	31	88.6	387	ADS24321	AdS24321 Bacterial
3	31	88.6	505	ABB58292	Abb58292 Drosophil
4	31	88.6	517	ADM06141	Adm06141 Human pro
5	31	88.6	654	AAE16784	Aae16784 Human tra
6	31	88.6	655	ABJ37897	Abj37897 NOVX prot
7	31	88.6	686	AAE60336	Aae60336 C1C-K1 pr
8	31	88.6	687	AYI13937	Ayi13937 Human CLC
9	30	85.7	8	AAE30029	Aae30029 Human hyp
10	30	85.7	572	AYH81595	Ayh81595 Streptoco
11	30	85.7	650	AAU37748	Aau37748 Streptoco
12	30	85.7	650	ABU01272	Abu01272 S. pneumo
13	30	85.7	650	ABU45986	Abu45986 Protein e
14	30	85.7	650	ADK48559	Adk48559 Streptoco
15	30	85.7	650	ADM92138	Adm92138 S pneumon
16	30	85.7	654	ADR94315	Adr94315 Novel S.
17	29	82.9	99	ADD19323	Add19323 Human sec
18	29	82.9	120	ABU05818	Abu05818 M. tuberc
19	29	82.9	120	ABP57468	Abp57468 Mycobacte
20	29	82.9	248	ADS23221	AdS23221 Bacterial
21	29	82.9	344	AAE23267	Aae23267 Arabidops
22	29	82.9	353	ABE58318	AbE58318 Drosophil
23	29	82.9	361	AAE23266	Aae23266 Arabidops
24	29	82.9	367	ADK24002	AdK24002 Bacterial
25	29	82.9	378	ABU20431	Abu20431 Protein e

26	29	82.9	436	2	AAV35033	Aay35033 Chlamydia
27	29	82.9	441	3	AAE23265	Aag23265 Arabidops
28	29	82.9	444	4	ABG05484	Abg05484 Novel hum
29	29	82.9	463	5	ABB49960	Abb49960 Listeria
30	29	82.9	493	4	AAE96119	Aae96119 Putative
31	29	82.9	578	8	ADQ65875	Adq65875 Novel hum
32	29	82.9	609	8	ADQ66688	Adq66688 Novel hum
33	29	82.9	751	6	AAE35491	Aae35491 Streptomy
34	29	82.9	798	6	ABU17132	Abu17132 Protein e
35	29	82.9	814	6	ADA33086	Ada33086 Acinetoba
36	29	82.9	819	8	ADP99142	Adp99142 Human tra
37	29	82.9	908	3	AAE42511	Aae42511 Human ORF
38	29	82.9	1307	6	AAO31015	Aao31015 Human tra
39	29	82.9	1353	6	AAE29913	Aae29913 Human tra
40	29	82.9	1359	8	ADS75981	AdS75981 Human ATP
41	29	82.9	1381	8	ADP99172	Adp99172 Human tra
42	28	80.0	8	5	ABP54724	Abp54724 Hypoxia i
43	28	80.0	8	5	ABP54723	Abp54723 Hypoxia i
44	28	80.0	8	5	ABP54721	Abp54721 Hypoxia i
45	28	80.0	8	6	AAE30025	Aae30025 Human hyp
46	28	80.0	8	6	AAE30028	Aae30028 Human hyp
47	28	80.0	12	7	AAO23489	Aao23489 Murine HI
48	28	80.0	12	7	AAO23517	Aao23517 Murine HI
49	28	80.0	14	6	AAE30165	Aae30165 Peptide #
50	28	80.0	14	7	AAO23514	Aao23514 Murine HI
51	28	80.0	14	7	AAO23486	Aao23486 Murine HI
52	28	80.0	15	7	AAO23513	Aao23513 Murine HI
53	28	80.0	15	7	AAO23485	Aao23485 Murine HI
54	28	80.0	16	6	AAE30163	Aae30163 Peptide #
55	28	80.0	16	7	AAO23488	Aao23488 Murine HI
56	28	80.0	16	7	AAO23512	Aao23512 Murine HI
57	28	80.0	16	7	AAO23484	Aao23484 Murine HI
58	28	80.0	16	7	AAO23516	Aao23516 Murine HI
59	28	80.0	17	7	AAO23483	Aao23483 Murine HI
60	28	80.0	17	7	AAO23511	Aao23511 Murine HI
61	28	80.0	18	6	ABP57669	Abp57669 Hypoxia-i
62	28	80.0	18	7	AAO23482	Aao23482 Murine HI
63	28	80.0	18	7	AAO23510	Aao23510 Murine HI
64	28	80.0	19	4	AAE49912	Aae49912 Human/mur
65	28	80.0	19	6	AAE30166	Aae30166 Peptide #

ALIGNMENTS

RESULT 1	ABR82377	ABR82377 standard; peptide; 8 AA.
ID	ABR82377	
XX	ABR82377	
AC	ABR82377	
DT	06-NOV-2003 (first entry)	
DE	Hypoxia-inducible factor 1 (HIF-1) alpha peptide inhibitor.	
XX	HIF-1; hypoxia-inducible factor 1; HIF-1 alpha; ubiquitination; EPO; erythropoietin; vascular endothelial growth factor; VEGF; glycolytic; tranquilizer; vulnerary; cardiant; cerebroprotective; angiogenesis.	
OS	Synthetic.	
XX	Key	Location/Qualifiers
XX	Modified-site	4 /label= Hyp
FT		/note= "hydroxyproline"
XX		
XX	WO2003057820-A2.	
PN		
XX	17-JUL-2003.	
PD		
XX		
XX	04-OCT-2002; 2002WO-US031699.	
PF		
XX		
XX	21-DEC-2001; 2001US-0002361.	
PR		

XX PA (KIMB) KIMBERLY-CLARK WORLDWIDE INC.

XX PI Mcgrath K;

XX DR WPT; 2003-645988/61.

XX DR Novel peptide inhibitor of hypoxia-inducible factor 1 alpha

PT ubiqutination, and activator of vascular endothelial growth factor

PT transcription useful for treating tissue injuries including wounds,

PT surgical incisions.

XX Claim 3; Page 8; 37pp; English.

XX The invention relates to peptide inhibitors of hypoxia-inducible factor

CC (HIF-1) alpha ubiquitination. The peptide inhibitors thereby activate the

CC transcripion of erythropoietin (EPO), vascular endothelial growth factor

CC (VEGF) and certain glycolytic enzymes. The peptide inhibitors are useful

CC for treating tissue injuries including wounds, surgical incisions,

CC chronic wounds, heart disease and stroke. The present sequence represents

CC a specific example of HIF-1 alpha peptide inhibitor

XX Sequence 8 AA;

SQ

Query Match 100.0%; Score 35; DB 6; Length 8;

Best Local Similarity 87.5%; Pred. No. 1.8e+06;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8

Db 1 MLAPTIPM 8

RESULT 2

ADS24321

ID ADS24321 standard; protein; 387 AA.

XX AC ADS24321;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #13354.

XX KW Recombinant DNA construct; transformed plant; improved plant property;

XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

XX KW pathogen tolerance; pest tolerance; plant disease resistance;

XX KW cell cycle pathway modification; plant growth regulator;

XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;

XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

XX KW bacterial polypeptide.

XX OS Bacteria.

XX US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOV/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX DR New recombinant DNA construct comprising a promoter positioned to provide

XX PT for expression of a polynucleotide encoding a polypeptide from a

PT

PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 13354; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a

CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plants with

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomannan

CC production. This sequence represents a bacterial polypeptide used in the

CC scope of the invention. Note: The sequence data for this patent did not

CC form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 387 AA;

Query Match 88.6%; Score 31; DB 8; Length 387;

Best Local Similarity 75.0%; Pred. No. 2.8e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8

Db 97 MLTSTIPM 104

RESULT 3

ABB58292

ID ABB58292 standard; protein; 505 AA.

XX AC ABB58292;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 1668.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX DR N-PSDB; ABL02395.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

XX PS Disclosure; SEQ ID NO 1668; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent

XX CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

XX CC useful in developmental biology and in elucidating cell signalling and

XX CC cell-cell interactions in higher eukaryotes for the development of

XX CC insecticides, therapeutics and pharmaceutical drugs. The invention

XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-

XX CC ABB72072). The sequence data for this patent did not form part of the

XX CC printed specification, but was obtained in electronic format directly

XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 505 AA;

Query Match 88.6%; Score 31; DB 4; Length 505;

Best Local Similarity 75.0%; Pred. No. 3.8e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8

Db 317 MLAXTIPV 324

RESULT 4

ADM06141

ID ADM06141 standard; protein; 517 AA.

XX AC ADM06141;

XX DT 20-MAY-2004 (first entry)

XX DE Human protein of the invention SEQ ID NO:4826.

XX KW human; gene therapy; diagnostic marker; pharmaceutical.

XX OS Homo sapiens.

XX PN EP1347046-A1.

XX PD 24-SEP-2003.

XX PF 12-APR-2002; 2002EP-00008400.

XX PR 22-MAR-2002; 2002JP-00137785.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-723558/69.

DR N-PSDB; ADM03698.

XX PT New polynucleotides and polypeptides are useful in gene therapy, for

XX PT developing a diagnostic marker or medicines for regulating their

XX PT expression and activity, or as a target of gene therapy.

XX PS Claim 1; SEQ ID NO 4826; 305pp; English.

XX CC The invention relates to a novel human polynucleotide and the encoded

XX CC polypeptide. A polynucleotide of the invention may have a use in gene

XX CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful

XX CC as a primer for synthesizing the polynucleotide or as a probe for

XX CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are

XX CC useful in gene therapy, for developing a diagnostic marker or medicines

XX CC for regulating their expression and activity, or as a target of gene

XX CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides

XX CC are useful as pharmaceutical agents. The present sequence represents a

XX CC protein sequence of the invention.

XX SQ Sequence 517 AA;

Query Match 88.6%; Score 31; DB 7; Length 517;

Best Local Similarity 75.0%; Pred. No. 3.9e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8

Db 245 ILAXTIPM 252

RESULT 5

AAE16784

ID AAE16784 standard; protein; 654 AA.

XX AC AAE16784;

XX DT 09-APR-2002 (first entry)

XX DE Human transporter and ion channel-21 (TRICH-21) protein.

XX KW Human; transporter and ion channel-21; TRICH-21; neuroprotective; asthma;

XX KW nootropic; cytostatic; cardiovascular; immunosuppressive; cardiomyopathy;

XX KW antiinflammatory; protein therapy; akinesia; cystic fibrosis; leukaemia;

XX KW Bell's palsy; amyotrophic lateral sclerosis; Alzheimer's disease; cancer;

XX KW amnesia; dementia; myocarditis; Duchenne's muscular dystrophy; AIDS;

XX KW Acquired Immune Deficiency Syndrome; Addison's disease; allergy; angina;

XX KW cell proliferative disorder; psoriasis; cardiac disease; hypertension;

XX KW bradyarrhythmia; gene expression; drug screening.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 49..72

FT /label= Transmembrane_domain

FT Domain 67..484

FT /note= "Voltage gated chloride channel domain"

FT Domain 364..382

FT /label= Transmembrane_domain

XX PN WO200192304-A2.

XX PD 06-DEC-2001.

XX PF 25-MAY-2001; 2001WO-US017065.

XX PR 26-MAY-2000; 2000US-0208424P.

XX PR 01-JUN-2000; 2000US-0209001P.

XX PR 08-JUN-2000; 2000US-0210588P.

XX PR 16-JUN-2000; 2000US-0212335P.

XX PR 22-JUN-2000; 2000US-0213747P.

XX PR 29-JUN-2000; 2000US-0215391P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Thornton M, Wallia NK, Yue H, Nguyen DB, Lai P, Gandhi AR;

XX PI Triboulet CM, Yao MG, Ramkumar J, Au-Young J, Lu Y, Tang YT;

XX PI Azimzai Y, Bruns CM, Griffin JA, Yang J, Sanjanwala MS, Raumann BE;

XX PI Lee EA, Hafalia A, Baughn MR, Green BD, Khan FA, Kearney L;

XX PI Elliot VS, Seilhamer JJ, Policky JU, Borowsky ML, Burford N, Ding L;

XX PI Lu DAM, Hillman JL;

XX WPI; 2002-122055/16.

DR N-PSDB; RAD27274.

XX PT New human transporters and ion channels (TRICH) polypeptides useful for

XX PT diagnosing, treating or preventing disorders associated with aberrant

XX PT expression of TRICH.

XX PS Claim 1; Page 176-178; 210pp; English.

XX CC The invention relates to human transporters and ion channels (TRICH)

CC polypeptides and their cDNA molecules. The nucleic acid and polypeptide
 CC sequences are useful in the diagnosis, treatment, and prevention of
 CC disorders associated with transport (akinesia, cystic fibrosis, Bell's
 CC palsy, amyotrophic lateral sclerosis); neurological (Alzheimer's disease,
 CC amnesia, dementia); muscle (cardiomyopathy, myocardiitis, Duchenne's
 CC muscular dystrophy); immunological (AIDS, Addison's disease, allergies,
 CC asthma); cell proliferative disorders (cancers, leukaemia, psoriasis);
 CC cardiac disease (angina, hypertension, or bradyarrhythmia) and in the
 CC assessment of the effects of exogenous compounds on the expression of
 CC nucleic acid and amino acid sequences of transporters and ion channels.
 CC The polynucleotides may be used to detect and quantify gene expression in
 CC biopsied tissues in which TRICH expression may be correlated with a
 CC disease, to generate hybridization probes for mapping naturally occurring
 CC genomic sequence, and in drug screening. The present sequence is human
 CC TRICH-21 protein
 CC
 XX
 SQ Sequence 654 AA;

Query Match 88.6%; Score 31; DB 5; Length 654;
 Best Local Similarity 75.0%; Pred. No. 5e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 2; Mismatches 0;

QY 1 MLAXTIPM 8
 :||:||||
 Db 382 ILATIPM 389

RESULT 6
 ABJ37897
 ID ABJ37897 standard; protein; 655 AA.
 XX
 AC ABJ37897;
 XX
 DT 22-MAY-2003 (first entry)
 XX
 DE NOVX protein sequence SEQ ID No 40.
 XX
 KW Hepatotropic; immunosuppressive; cardiant; hypertensive; tranquilizer;
 KW vulnary; virucide; antibacterial; protozoicide; fungicide; nootropic;
 KW antiparasitic; neuroprotective; cerebroprotective; antiparkinsonian;
 KW anticonvulsant; antidiabetic; analgesic; dermatological; keratolytic;
 KW antiseborrheic; antirheumatic; antiarthritic; antinflammatory; anti-HIV;
 KW cytostatic; antiaesthetic; antipsoriatic; hypotensive; osteopathic;
 KW antitumor; anorectic; antidiabetic; antiallergic; haemostatic;
 KW neuroleptic; antidepressant; antifertility; NOVX; human disease;
 KW NOVX-associated disorder; trauma; viral; bacterial; fungal; protozoal;
 KW parasitic infection; Alzheimer's disease; stroke; forensic biology;
 KW immunogen; non-human transgenic animal; gene therapy.
 XX
 OS Unidentified.
 XX
 PN WO200281517-A2.
 XX
 PD 17-OCT-2002.
 XX
 PE 22-JAN-2002; 2002WO-US002064.
 XX
 PR 19-JAN-2001; 2001US-0262892P.
 PR 23-JAN-2001; 2001US-0263598P.
 PR 24-JAN-2001; 2001US-0263799P.
 PR 25-JAN-2001; 2001US-0264117P.
 PR 26-JAN-2001; 2001US-0264139P.
 PR 28-JAN-2001; 2001US-0264478P.
 PR 30-JAN-2001; 2001US-0263351P.
 PR 02-MAR-2001; 2001US-0272870P.
 PR 14-MAR-2001; 2001US-0275927P.
 PR 15-MAR-2001; 2001US-0276449P.
 PR 20-MAR-2001; 2001US-0277358P.
 PR 23-MAR-2001; 2001US-0278151P.
 PR 29-MAR-2001; 2001US-0279857P.
 PR 20-APR-2001; 2001US-0285140P.
 PR 20-APR-2001; 2001US-0285141P.

30-APR-2001; 2001US-0287484P.
 17-MAY-2001; 2001US-0291701P.
 PR 08-JUN-2001; 2001US-0296960P.
 PR 10-JUL-2001; 2001US-0304353P.
 PR 12-JUL-2001; 2001US-0304353P.
 PR 09-AUG-2001; 2001US-0304886P.
 PR 13-AUG-2001; 2001US-0311289P.
 PR 16-AUG-2001; 2001US-0311975P.
 PR 18-OCT-2001; 2001US-0312937P.
 PR 29-NOV-2001; 2001US-0330227P.
 PR 29-NOV-2001; 2001US-0334198P.
 (CURA-) CURAGEN CORP.
 Decristofaro MF, Padigaru M, Miller C, Tchernev V, Zhong H;
 PI Zhong M, Anderson D, Ballinger R, Gerlach V, Spytek KA, Rastelli L;
 PI Kekuda R, Guo X, Zerhusen B, Andrew D, Mezes P, Patturajan M;
 PI Burgess CE, Eisen A, Wolenc A, Baumgartner J, Shinkets RA, Gusev V;
 PI Vernet CAM, Taupier RJ, Pena C, Shenoy S, Li L, Casman S, Boldog F;
 PI Fernandes E, Smithson G, Malyankar U, Tallon B, Liu X;
 WPI; 2003-058504/05.
 N-PSDB; AB733362.
 New polypeptides, designated as NOVX, useful for diagnosing and treating
 infections, neurological diseases, cancer, allergy, and bone,
 immunological, skin, renal, brain, muscle and autoimmune disorders.
 Claim 1; Page 111; 672pp; English.

The invention relates to a novel isolated polypeptide, designated NOVX
 (NOV1 - 33), consisting of a mature form of one of 61 sequences, given in
 the specification, or its variant, where amino acid residue(s) in the
 variant differ from the mature form, provided that the variant differs in
 not more than 15 % of the amino acids from the sequence of the mature
 form. The NOVX polypeptides, nucleic acids encoding the polypeptides, and
 an antibody to the polypeptides, are useful for treating or preventing a
 NOVX-associated disorder in humans and for treating a syndrome associated
 with a human disease (NOVX-associated disorder). NOVX polypeptides and
 the encoding nucleic acids, are useful for determining the presence of or
 predisposition to a disease associated with altered levels of NOVX
 polypeptide and polynucleotide, by measuring the level of polypeptide
 expression or the amount of nucleic acid from a mammal and comparing it
 with another mammal not having or not predisposed to the disease. NOVX
 polypeptide is also useful for identifying an agent that binds to NOVX
 and a cell expressing NOVX is useful for identifying an agent that
 modulates the expression or activity of NOVX. The antibodies and a
 polypeptide having 95 % sequence identity to NOVX polypeptides are also
 useful for treating a pathological state in a mammal. The antibodies are also
 useful for determining the presence or amount of NOVX in a sample. NOVX
 polypeptides, polynucleotides and antibodies specific for the
 polypeptides are useful for treating or preventing disorders or syndromes
 including trauma, viral, bacterial, fungal, protozoal, and parasitic
 infections. They can also treat disorders such as e.g., Alzheimer's
 disease or a stroke. The NOVX encoding nucleic acids are useful for
 expressing the NOVX proteins, to detect NOVX mRNA, or a genetic lesion in
 a NOVX gene and to modulate NOVX activity. NOVX sequences are also useful
 for identifying a cell or tissue type in a biological sample, to amplify
 DNA sequences from very small biological samples such as tissues e.g.
 hair or skin or body fluids in forensic biology and as primers and probes
 for use in identifying and/or cloning NOVX homologues in other cell
 types. The NOVX proteins are useful as an immunogen to generate
 antibodies which are useful for diagnostically monitoring protein levels
 and modulating NOVX activity. Cells comprising NOVX nucleic acids are
 useful for producing non-human transgenic animals which are useful for
 studying the function and/or activity of NOVX protein and for identifying
 and/or evaluating modulators of NOVX protein activity. The NOVX nucleic
 acids can be used in gene therapy. This sequence represents a NOVX
 protein of the invention

Sequence 655 AA;

Query Match 88.6%; Score 31; DB 6; Length 655;


```

Best Local Similarity 75.0%; Pred. No. 5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 382 ILATTIPM 389
      :||:||||

RESULT 7
ID AAR60336 standard; protein; 686 AA.
AC AAR60336;
XX
XX
DT 16-OCT-2003 (revised)
DT 27-FEB-1995 (first entry)
XX
XX
DE C1C-K1 protein.
XX
KW C1C-K1; kidney; Henle's loop; probe; detection; chloride; diagnosis.
XX
OS Rattus rattus; - Sprague-Dawley.
XX
PN JP06165680-A.
XX
XX 14-JUN-1994.
XX
PF 01-DEC-1992; 92JP-00343609.
XX
PR 01-DEC-1992; 92JP-00343609.
XX
PA (MARU/) MARUMO F.
XX
DR WPI; 1994-230226/28.
DR N-PSDB; AAQ70421.
XX
PT C1C-K1 cDNA probe specific for kidney Henle's loop kicked ascending limb
PT maldistributed chloride channel RNA - useful in diagnosis of diseases
PT caused by abnormal chloride transition.
XX
PS Claim 8; Page 7-10; 14pp; Japanese.
XX
XX The C1C-K1 probes (AAQ70422-23) are used in the detection of C1C-K1 DNA.
CC Detection is useful in diagnosis of diseases caused by abnormal
CC transition of chloride. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 686 AA;

Query Match 88.6%; Score 31; DB 2; Length 686;
Best Local Similarity 75.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 414 ILATTIPM 421
      :||:||||

RESULT 8
ID AAY13937 standard; protein; 687 AA.
XX
XX AAY13937;
AC
XX
DT 14-JUL-1999 (first entry)
XX
DE Human CLCNKB protein.
XX
KW CLCNKB gene; human; PCR primer; renal chloride ion transporter;
KW diagnosis; renal ion transport defect; Bartter's syndrome; gene therapy;
KW Gitelman's syndrome; hypotensive condition; hypertensive condition;
KW congestive heart failure.
XX
OS Homo sapiens.

Query Match 88.6%; Score 31; DB 2; Length 687;
Best Local Similarity 75.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 414 ILATTIPM 421
      :||:||||

RESULT 9
ID AAE30029 standard; peptide; 8 AA.
XX
XX AAE30029;
AC
XX
DT 24-FEB-2003 (first entry)
XX
DE Human hypoxia-inducible factor 1 (HIF-1) alpha derived peptide #5.
XX
KW Entity localisation; light-generating fusion protein; LGP; diabetes;
KW hypoxic tissue imaging; detection; cancer; hypoxia; heart disease;
KW stroke; human; hypoxia-inducible factor 1; HIF-1 alpha.
XX
OS Homo sapiens.
XX
PN WO200275278-A2.
XX
PD 26-SEP-2002.
XX
PF 20-MAR-2002; 2002WO-US008864.
XX
PR 20-MAR-2001; 2001US-0277425P.
PR 20-MAR-2001; 2001US-0277431P.
PR 20-MAR-2001; 2001US-0277440P.

```

```

XX WO9916909-A1.
FN
XX
XX 08-APR-1999.
XX
XX 01-OCT-1998; 98WO-US020777.
PF
XX
PR 01-OCT-1997; 97US-0060219P.
XX
XX (UYYA ) UNIV YALE.
PA
XX
XX Lifton RP, Simon DB;
PI
XX
XX WPI; 1999-263707/22.
DR
XX N-PSDB; AAX36795.
DR
XX
XX Diagnosing a pathological condition associated with a defect in renal ion
XX transport.
XX
XX Disclosure; Fig 10b; 80pp; English.
XX
XX This sequence represents the human renal chloride ion transporter,
XX CLCNKB, protein. The invention relates to a method for diagnosing a
XX pathological condition associated with a defect in renal ion transport in
XX a subject, by determining whether a renal chloride channel (CLCNKB) gene
XX is present or mutated in a patient. The nature of the mutation
XX contributes to making a differential diagnosis between Bartter's syndrome
XX Type III and Gitelman's syndrome, other hypo- and hyper-tensive
XX conditions and Bartter's syndrome type III (heterozygous alteration).
XX allow diagnosis of Bartter's syndrome type III (heterozygous alteration)
XX and a carrier of Bartter's syndrome type III (heterozygous alteration).
XX Modulators of CLCNKB can be used for treatment of hypertension and
XX congestive heart failure. The CLCNKB gene and protein can also serve as
XX targets for gene therapy. Knowledge of the mutations in CLCNKB proteins
XX and genes will result in easier classification of Gitelman's Syndrome and
XX Bartter's Syndrome, without having to rely solely on clinical features
XX
XX Sequence 687 AA;

Query Match 88.6%; Score 31; DB 2; Length 687;
Best Local Similarity 75.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 414 ILATTIPM 421
      :||:||||

```

XX	Haselbeck R,	Ohlsen KL,	Zyskind JW,	wall D,	Trawick JD,	Carr GJ;
PI	Yamamoto RT,	Xu HH;				
PI						

XX WPI; 2001-611495/70.
 DR N-PSDB; AAS55607.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 PT
 XX
 XX Example 3; SEQ ID NO 13341; 511pp; English.
 PS
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 650 AA;
 SQ

Query Match 85.7%; Score 30; DB 4; Length 650;
 Best Local Similarity 75.0%; Pred. NO. 8.4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MLAXTIEM 8
 ||||: ||||
 DB 458 MLAVNIEM 465

RESULT 12
 ABU01272
 ID ABU01272 standard; protein; 650 AA.
 XX
 AC ABU01272;
 XX
 XX 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX
 DE S. pneumoniae type 4 strain protein from coding region #847.
 XX
 KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae; type 4 strain.
 XX
 XX WO200277021-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 27-MAR-2002; 2002WO-IB002163.
 PF
 XX 27-MAR-2001; 2001GB-00007658.
 PR
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 XX Masignani V, Tetelin H, Fraser C;
 PI
 XX WPI; 2003-040579/03.
 DR
 DR N-PSDB; ABX06559.
 XX
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT

PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 PT ear infection.
 XX
 XX Claim 1; SEQ ID NO 1694; 56pp; English.
 PS
 XX The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC AB556454. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC sequence contained within a Streptococcus nucleic acid sequence, where
 CC the first primer is substantially complementary to the target sequence
 CC and the second primer is substantially complementary to the complement of
 CC the target sequence, and where the parts of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound with the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 650 AA;
 OY 1 MLAXTIEM 8
 ||||: ||||
 DB 458 MLAVNIEM 465

Query Match 85.7%; Score 30; DB 6; Length 650;
 Best Local Similarity 75.0%; Pred. NO. 8.4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MLAXTIEM 8
 ||||: ||||
 DB 458 MLAVNIEM 465

RESULT 13
 ABU45986
 ID ABU45986 standard; protein; 650 AA.
 XX
 AC ABU45986;
 XX
 XX 19-JUN-2003 (first entry)
 DT
 XX
 XX Protein encoded by Prokaryotic essential gene #31513.
 DE
 XX
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW
 XX Streptococcus pneumoniae.
 OS
 XX WO200277183-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002WO-US009107.
 PF
 XX 21-MAR-2001; 2001US-00815242.
 PR
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 PR

XX PA (ELIT-) ELITRA PHARM INC. ,
XX PI Wang L, Zamudio C, Malone G, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Tawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WP1; 2003-029926/02.
DR N-PSDB; ACA49856.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 73910; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 650 AA;
Query Match 85.7%; Score 30; DB 6; Length 650;
Best Local Similarity 75.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAXTIPM 8
Db 458 MLAVNIPM 465
||||: |||
RESULT 14
ADK48559
ID ADK48559 standard; protein; 650 AA.
XX AC ADK48559;
XX 20-MAY-2004 (first entry)
XX Streptococcus pneumoniae protein, Seq ID NO 5074.
XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX Streptococcus pneumoniae.
XX US6699703-B1.
XX PN

XX PD 02-MAR-2004.
XX 26-MAY-2000; 2000US-00583110.
XX 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX N-PSDB; ADK45898.
XX WPI; 2004-212399/20.
XX New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.
XX Disclosure; SEQ ID NO 5074; 301pp; English.
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as *S. pneumoniae* infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX SQ Sequence 650 AA;
Query Match 85.7%; Score 30; DB 8; Length 650;
Best Local Similarity 75.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAXTIPM 8
Db 458 MLAVNIPM 465
||||: |||
RESULT 15
ADM92138
ID ADM92138 standard; protein; 650 AA.
XX AC ADM92138;
XX 03-JUN-2004 (first entry)
XX S pneumoniae antigenic protein sequence SeqID335.
XX antibacterial; gene therapy; Streptococcus pneumoniae infection;
XX antigenic.
XX Streptococcus pneumoniae.
XX WO2004020609-A2.
XX 11-MAR-2004.
XX 02-SEP-2003; 2003WO-US027401.
XX 30-AUG-2002; 2002US-0407082P.
XX (TUFT) UNIV TUFTS.
XX Camilli A, Hava DL;
XX WPI; 2004-239189/22.
XX N-PSDB; ADM91901.
XX DR

XX New Streptococcus pneumoniae nucleic acid molecules, useful for
PT diagnosing, treating and preventing active infections of Streptococcus
PT pneumoniae.
XX
XX
XX Claim 27; SEQ ID NO 335; 123pp; English.
PS
PS This invention relates to novel isolated Streptococcus pneumoniae nucleic
CC acid molecules and the antigenic polypeptides encoded by them. The
CC invention may be useful for the production of compounds with an
CC antibacterial activity or for gene therapy. The nucleic acid molecules,
CC compositions and methods disclosed are useful for treating Streptococcus
CC pneumoniae infection. The present sequence is that of an S pneumoniae
CC protein of the invention.
XX
XX SQ Sequence 650 AA;
Query Match 85.7%; Score 30; DB 8; Length 650;
Best Local Similarity 75.0%; Pred. No. 8.4e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAXTIPM 8
|||: |||
Db 458 MLAVNIPM 465
RESULT 16
ADR94315
ID ADR94315 standard; protein; 654 AA.
XX
XX ADR94315;
AC
XX
DT 16-DEC-2004 (first entry)
DE
DE Novel S. pneumoniae protein sequence, SEQ ID 2950.
XX
XX Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
KW bacterial infection.
XX
XX Streptococcus pneumoniae.
XX
XX US6800744-B1.
XX
XX 05-OCT-2004.
XX
XX 30-JUN-1998; 98US-00107433.
XX
XX 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
PI
XX WPI; 2004-697205/68.
DR N-PSDB; ADR91712.
XX
XX New isolated nucleic acid encoding a Streptococcus pneumoniae
PT polypeptide, useful for diagnosing, preventing and/or treating
PT pathological conditions resulting from the bacterial infection.
XX
XX Disclosure; SEQ ID NO 2950; 151pp; English.
PS
PS The invention relates to an isolated nucleic acid comprising a sequence
CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its
CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,
CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or
CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
CC hybridisable under high stringency conditions to the nucleotide sequences.
CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.

CC Also included are a recombinant expression vector comprising the isolated
CC nucleic acid cited above operably linked to a transcription regulatory
CC element, a cell comprising the recombinant expression vector and a probe
CC comprising at least 20 consecutive nucleotides of the nucleotide
CC sequences as cited above. The methods and compositions of the present
CC invention are useful for the diagnosis, prevention and/or treatment of
CC pathological conditions resulting from bacterial infection by
CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and
CC otitis media. The present sequence is one of the 2603 disclosed S.
CC pneumoniae protein sequences. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.
XX
XX SQ Sequence 654 AA;
Query Match 85.7%; Score 30; DB 8; Length 654;
Best Local Similarity 75.0%; Pred. No. 8.5e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAXTIPM 8
|||: |||
Db 462 MLAVNIPM 469
RESULT 17
ADD19323
ID ADD19323 standard; protein; 99 AA.
XX
XX ADD19323;
AC
XX
DT 15-JAN-2004 (first entry)
DE
DE Human secreted protein from gene 29 #2.
XX
XX human secreted protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW neurotic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; human.
XX
XX Homo sapiens.
XX
XX WO2003052377-A2.
PN
XX 26-JUN-2003.
PD
XX 06-NOV-2002; 2002WO-US035606.
PP
XX 07-NOV-2001; 2001US-0331046P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-533050/50.
XX
XX N-PSDB; ADD19248.
DR
XX
XX New isolated nucleic acids encoding signal transduction pathway component
PT polypeptides, useful for diagnosing, treating, and/or preventing
PT disorders, such as cancer, infections, cardiovascular and inflammatory
PT diseases.
XX
XX Claim 11; SEQ ID NO 150; 554pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human secreted protein, representing one of 85 novel genes.
CC Also included are recombinant vectors, host cells (expressing the
CC protein), the secreted proteins (including their fragments, epitopes and

XX
I.7XX
I.7

CC and immunostimulant activities, and can be used in vaccines and gene
 CC therapy. (i) and (ii) can be used for the manufacture of a medicament for
 CC treating or preventing a mycobacterial infection. They can also be used
 CC for the manufacture of a diagnostic reagent for identifying a
 CC mycobacterial infection

XX Sequence 120 AA;

QQ Sequence 120 AA; Query Match 82.9%; Score 29; DB 6; Length 120;

Best Local Similarity 62.5%; Pred. No. 2.1e+02;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8

DB 94 LLAGTIPL 101

RESULT 20

ID ADS23221 standard; protein; 248 AA.

XX AC ADS23221;

DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #12254.

KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 12254; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 248 AA;

Query Match 82.9%; Score 29; DB 8; Length 248;

Best Local Similarity 71.4%; Pred. No. 4.7e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8

DB 204 LALTVPM 210

RESULT 21

AAG23267

ID AAG23267 standard; protein; 344 AA.

XX AC AAG23267;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 26510.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 23-APR-1999; 99US-0130891P.

XX PR 28-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 04-MAY-1999; 99US-0132484P.

XX PR 05-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 07-MAY-1999; 99US-0132863P.

XX PR 11-MAY-1999; 99US-0132867P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134221P.

XX PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 08-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 18-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151910P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.

PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.9%; Score 29; DB 3; Length 344;
Best Local Similarity 71.4%; Pred. No. 6.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
||:|:|
DB 179 LAYTVM 185

RESULT 22
ABBS8318
ID ABBS8318 standard; protein; 353 AA.

XX AC ABBS8318;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 1746.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL02421.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.

XX PS Disclosure; SEQ ID NO 1746; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 353 AA;

Query Match 82.9%; Score 29; DB 4; Length 353;
Best Local Similarity 71.4%; Pred. No. 7.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
||:|:|

Db 177 MLACTVP 183

RESULT 23

AAG23266

ID AAG23266 standard; protein; 361 AA.

XX AC AAG23266;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 26509.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 23-MAR-1999; 99US-0123548P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 01-APR-1999; 99US-0126785P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 28-APR-1999; 99US-0130891P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 30-APR-1999; 99US-0132407P.

XX PR 04-MAY-1999; 99US-0132484P.

XX PR 05-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 07-MAY-1999; 99US-0132863P.

XX PR 11-MAY-1999; 99US-0134256P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134221P.

XX PR 18-MAY-1999; 99US-0134370P.

XX PR 19-MAY-1999; 99US-0134941P.

XX PR 20-MAY-1999; 99US-0135124P.

XX PR 21-MAY-1999; 99US-0135353P.

XX PR 24-MAY-1999; 99US-0135629P.

XX PR 25-MAY-1999; 99US-0136021P.

XX PR 27-MAY-1999; 99US-0136392P.

XX PR 28-MAY-1999; 99US-0136782P.

XX PR 01-JUN-1999; 99US-0137222P.

XX PR 03-JUN-1999; 99US-0137528P.

XX PR 04-JUN-1999; 99US-0137502P.

XX PR 07-JUN-1999; 99US-0137724P.

XX PR 08-JUN-1999; 99US-0138094P.

XX PR 10-JUN-1999; 99US-0138540P.

XX PR 10-JUN-1999; 99US-0138847P.

XX PR 14-JUN-1999; 99US-0139119P.

XX PR 16-JUN-1999; 99US-0139452P.

XX PR 16-JUN-1999; 99US-0139453P.

XX PR 17-JUN-1999; 99US-0139492P.

XX PR 18-JUN-1999; 99US-0139454P.

XX PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143622P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.

PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151303P.
PR 30-AUG-1999; 99US-0151438P.
PR 31-AUG-1999; 99US-0151930P.
PR 01-SEP-1999; 99US-0152363P.
PR 07-SEP-1999; 99US-0153070P.
PR 10-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 15-SEP-1999; 99US-0154039P.
PR 16-SEP-1999; 99US-0154779P.
PR 20-SEP-1999; 99US-0155139P.
PR 22-SEP-1999; 99US-0155486P.
PR 23-SEP-1999; 99US-0155659P.
PR 24-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 29-SEP-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-0157753P.
PR 05-OCT-1999; 99US-0157865P.
PR 06-OCT-1999; 99US-0158029P.
PR 07-OCT-1999; 99US-0158232P.
PR 08-OCT-1999; 99US-0158369P.
PR 12-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.9%; Score 29; DB 3; Length 361;
Best Local Similarity 71.4%; Pred. No. 7.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 LAXTIPM 8
||:|:
Db 196 LAYTVP 202

RESULT 24
ADS24002
ID ADS24002 standard; protein; 367 AA.
XX

AC ADS24002;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #13035.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
OS US2003233675-A1.
XX
PN 18-DEC-2003.
XX
PD 20-FEB-2003; 2003US-00369493.
PF 21-FEB-2002; 2002US-0360039P.
PR (CAO Y) CAO Y.
XX (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 13035; 122pp; English.
PS
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 367 AA;
Query Match 82.9%; Score 29; DB 8; Length 367;
Best Local Similarity 71.4%; Pred. No. 7.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLAUTIP 7
[:::]

Db 293 MIACTVP 299
RESULT 25
ABU20431
ID ABU20431 standard; protein; 378 AA.
XX
AC ABU20431;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #5958.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bacteroides fragilis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA24301.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 48355; 1766pp; English.
PS
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 378 AA;

SQ Query Match 82.9%; Score 29; DB 6; Length 378;

Best Local Similarity 62.5%; Pred. No. 7.6e+02;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8

Db 247 LLSITIPM 254

RESULT 26

AAAY35033
ID AAY35033 standard; protein; 436 AA.

XX AC AAY35033;

XX DT 17-OCT-2003 (revised)
DT 13-SEP-1999 (first entry)

XX DE Chlamydia pneumoniae transmembrane protein sequence.

XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
KW neutralising epitope.

XX OS Chlamydothila pneumoniae.

XX PN WO927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB001890.

XX PR 21-NOV-1997; 97FR-00014673.

XX PR 04-NOV-1998; 98US-0107078P.

XX PA (GEST) GENSET.

XX PI Griffais R;

XX WPI; 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae.

XX PS Page 926-927; Disclosure; 1912pp; English.

XX CC AAY34584-Y35879 represent the proteins encoded by all the open reading
frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
XX CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
and is thought to be a contributing factor in heart disease, sarcoidosis,
XX CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
XX CC polypeptides encoded by the open reading frames of the C. pneumoniae
XX CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
XX CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
XX CC be used as immunogenic compositions, especially where the vector directs
XX CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
XX CC -OCT-2003 to standardise OS field)

XX SQ Sequence 436 AA;

Query Match 82.9%; Score 29; DB 2; Length 436;

Best Local Similarity 71.4%; Pred. No. 9e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8

Db 148 LAATVPM 154

RESULT 27.

AAG23265
ID AAG23265 standard; protein; 441 AA.
XX AC AAG23265;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 26508.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 23-MAR-1999; 99US-0123548P.
XX PR 25-MAR-1999; 99US-0125788P.
XX PR 29-MAR-1999; 99US-0126264P.
XX PR 01-APR-1999; 99US-0126785P.
XX PR 06-APR-1999; 99US-0127462P.
XX PR 08-APR-1999; 99US-0128234P.
XX PR 16-APR-1999; 99US-0128714P.
XX PR 19-APR-1999; 99US-0129845P.
XX PR 21-APR-1999; 99US-0130077P.
XX PR 23-APR-1999; 99US-0130449P.
XX PR 28-APR-1999; 99US-0130510P.
XX PR 30-APR-1999; 99US-0130891P.
XX PR 04-MAY-1999; 99US-0131449P.
XX PR 05-MAY-1999; 99US-0132048P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 11-MAY-1999; 99US-0132863P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 18-MAY-1999; 99US-0134370P.
XX PR 19-MAY-1999; 99US-0134768P.
XX PR 20-MAY-1999; 99US-0134941P.
XX PR 21-MAY-1999; 99US-0135124P.
XX PR 24-MAY-1999; 99US-0135353P.
XX PR 25-MAY-1999; 99US-0135629P.
XX PR 28-MAY-1999; 99US-0136021P.
XX PR 01-JUN-1999; 99US-0136392P.
XX PR 03-JUN-1999; 99US-0136782P.
XX PR 04-JUN-1999; 99US-0137222P.
XX PR 07-JUN-1999; 99US-0137528P.
XX PR 08-JUN-1999; 99US-0137724P.
XX PR 10-JUN-1999; 99US-0138094P.
XX PR 10-JUN-1999; 99US-0138540P.
XX PR 14-JUN-1999; 99US-0138847P.
XX PR 16-JUN-1999; 99US-0139119P.
XX PR 17-JUN-1999; 99US-0139452P.
XX PR 18-JUN-1999; 99US-0139453P.
XX PR 18-JUN-1999; 99US-0139454P.
XX PR 18-JUN-1999; 99US-0139455P.
XX PR 18-JUN-1999; 99US-0139456P.
XX PR 18-JUN-1999; 99US-0139457P.
XX PR 18-JUN-1999; 99US-0139458P.
XX PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.

PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 82.9%; Score 29; DB 3; Length 441;
Best Local Similarity 71.4%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTPM 8
||:|:|
Db 276 LAYTVP 282

RESULT 28
ABG05484
ID ABG05484 standard; protein; 444 AA.
XX
AC ABG05484;
XX
DT 13-FEB-2002 (first entry)
XX

DE Novel human diagnostic protein #5475.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS69671.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 35843; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 444 AA;
 Query Match 82.9%; Score 29; DB 4; Length 444;
 Best Local Similarity 62.5%; Pred. No. 9.2e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLAXTIPM 8
 Db 381 LLAVTIPL 388
 RESULT 29
 ABB49960
 ID ABB49960 standard; protein; 463 AA.
 XX ABB49960;
 AC ABB49960;
 XX 05-FEB-2002 (first entry)
 DT Listeria monocytogenes protein #2664.
 XX

KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 OS Listeria monocytogenes.
 XX WO200177335-A2.
 XX 18-OCT-2001.
 XX 11-APR-2001; 2001WO-FR001118.
 XX 11-APR-2000; 2000FR-00004629.
 XX (INSP) INST PASTEUR.
 XX Buchrieser C, Frangeul L, Couve E, Rueniok C, Fsihi H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst P, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Rierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX WPI; 2002-010914/01.
 DR Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 XX and prevention of Listeria and related bacterial infections, and related
 PT polypeptides.
 XX Claim 6; SEQ ID NO 2665; 192pp; French.
 XX The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 463 AA;
 Query Match 82.9%; Score 29; DB 5; Length 463;
 Best Local Similarity 50.0%; Pred. No. 9.6e+02;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLAXTIPM 8
 Db 365 LIASTVPM 372
 RESULT 30
 AAB96119
 ID AAB96119 standard; protein; 493 AA.
 XX AAB96119;
 AC AAB96119;
 XX 29-OCT-2001 (first entry)
 DT Putative P. abyssi NAD-dependent K+ and Na+ uptake protein.
 XX Hyperthermophilic archaeon; hyperthermophilic protein.
 XX

CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a protein
 CC sequence of the invention.

XX SQ Sequence 609 AA;
 Query Match 82.9%; Score 29; DB 8; Length 609;
 Best Local Similarity 62.5%; Pred. No. 1.3e+03;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLAXTIPM 8
 Db 184 LLSVTIPM 191
 :|:|:|

RESULT 33
 AAE35491
 ID AAE35491 standard; protein; 751 AA.
 AC AAE35491;
 DT 17-JUN-2003 (first entry)
 XX Streptomyces platensis subspecies roseaceus dorrigocin ORF8 protein.
 DE Polyketide biosynthesis; dorrigocin; DORR; lactimidomycin; LACT.
 KW Streptomyces platensis.
 OS Key Location/Qualifiers
 XX 1..277
 FH Domain /note= "Acyl transferase (AT) domain"
 FT 302..637
 FT Domain /note= "Oxidoreductase (Ox) domain"
 FT

XX WO200288176-A2.
 PN 07-NOV-2002.
 XX 26-APR-2002; 2002WO-CA000591.
 XX 26-APR-2001; 2001US-0286346P.
 XX (ECOP-) ECOPIA BIOSCIENCES INC.
 XX Farnet CM, Zazopoulos E, Staffa A, Yang X;
 PI WPI; 2003-201222/19.
 DR N-PSDB; AAD54225.

XX Novel isolated or purified polypeptide involved in biosynthesis of
 PT polyketide dorrigocin or polyketide lactimidomycin, useful for preparing
 PT dorrigocin or lactimidomycin.
 XX
 PS Claim 13; Page 191-193; 312pp; English.
 CC The invention relates to novel proteins involved in the biosynthesis of
 CC polyketide dorrigocin (DORR) or lactimidomycin (LACT) biosynthesis by
 CC microorganisms. Sequences of the invention allow direct manipulation of
 CC dorrigocin, lactimidomycin and related chemical structures via chemical
 CC engineering of the enzymes involved in the biosynthesis of dorrigocin and
 CC lactimidomycin. They are useful for introducing chemical handles into
 CC normally inert positions that permit subsequent chemical modifications
 CC and facilitate the development of polyketides. The genes and proteins of
 CC the invention can also be used to generate a focused library of analogues
 CC around a polyketide lead candidate to fine-tune the compound for optimal
 CC properties. They are useful for generating antibodies specific for the
 CC polyketide biosynthesis. The present sequence is Streptomyces platensis
 CC subspecies roseaceus DORR ORF8 protein

XX SQ Sequence 751 AA;
 Query Match 82.9%; Score 29; DB 6; Length 751;
 Best Local Similarity 71.4%; Pred. No. 1.7e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LAXTIPM 8
 Db 470 LARTVPM 476
 :|:|:|

RESULT 34
 ABU17132
 ID ABU17132 standard; protein; 798 AA.

XX AC ABU17132;
 XX DT 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #2659.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Acinetobacter baumannii.
 OS WO200277183-A2.
 PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA21002.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 45056; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 798 AA;

Query Match 82.9%; Score 29; DB 6; Length 798;

Best Local Similarity 71.4%; Pred. No. 1.8e+03;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MLAXTIP 7
 Db 730 MLAHTVP 736

RESULT 35

ADA33086
 ID ADA33086 standard; protein; 814 AA.

AC ADA33086;

XX 20-NOV-2003 (first entry)

DE Acinetobacter baumannii protein #247.

XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.

XX Acinetobacter baumannii.

OS US6562958-B1.

XX 13-MAY-2003.

PF 04-JUN-1999; 99US-00328352.

PR 09-JUN-1998; 98US-0088701P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Breton G, Bush D;

DR WPI; 2003-576092/54.

DR N-PSDB; ADA28960.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.

PS Example; SEQ ID NO 4373; 328pp; English.

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.

XX Sequence 814 AA;

Query Match 82.9%; Score 29; DB 6; Length 814;

Best Local Similarity 71.4%; Pred. No. 1.8e+03;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MLAXTIP 7
 Db 746 MLAHTVP 752

RESULT 36

ADP99142
 ID ADP99142 standard; protein; 819 AA.

AC ADP99142;

XX 26-AUG-2004 (first entry)

DE Human transporter and ion channel (TRICH) protein - SEQ ID 7.

XX transporter; ion channel; TRICH; neuroprotective; respiratory; muscular;
 KW antiparkinsonian; anticonvulsant; nootropic; neuroprotective;
 KW antiinflammatory; endocrine; anti-HIV; antianaemic; immunosuppressive;
 KW nephrotropic; hepatotropic; antipsoriatic; transport disorder;
 KW neurological; muscle; immunological; cell proliferative; transgenic;
 KW gene therapy; human.

XX Homo sapiens.

OS WO2004048599-A2.

XX 10-JUN-2004.

XX 20-NOV-2003; 2003WO-US037572.

XX 22-NOV-2002; 2002US-0428508P.

PR 06-DEC-2002; 2002US-0431635P.

PR 13-DEC-2002; 2002US-0433329P.

PR 24-JAN-2003; 2003US-0424233P.

PR 27-FEB-2003; 2003US-0451298P.

XX (INCY-) INCYTE CORP.

XX Tran UK, Jin P, Elliott VS, Yue H, Chawla NK, Yao MG;
 PI Raumann BE, Khan FA, Arvizu CS, Warren BA, Griffin JA, Ison CH;
 PI Hafalia AJA, Lal PG, Lee SY, Kable AB, Khare R, Swarnakar A;
 PI Becha SD, Marquis JP, Favero KD, Wang JT, Richardson TW, Bhatia UG;
 PI Burrill JD, Lee S, Blake JJ, Ho A, Zheng W, Nakamura Li;

XX WPI; 2004-441181/41.

DR N-PSDB; ADP99191.

XX New isolated human transporters and ion channels, TRICH 1-49, useful for
 PT diagnosing, treating, and preventing Duchenne's muscular dystrophy,
 PT polymyositis, psoriasis, Alzheimer's disease, and Cushing's syndrome.

PS Claim 1; SEQ ID NO 7; 292pp; English.

XX The invention relates to a novel isolated transporter and ion channel
 CC (TRICH) polypeptide comprising a polypeptide sequence as given in
 CC specification, or its biologically active or immunogenic fragment. The
 CC polypeptide of the invention demonstrates neuroprotective, respiratory,
 CC muscular, antiparkinsonian, anticonvulsant, nootropic, neuroprotective,
 CC antiinflammatory, endocrine, anti-HIV, antianaemic, immunosuppressive,
 CC nephrotropic, hepatotropic and antipsoriatic activities. The polypeptide
 CC may be useful in the diagnosis, treatment and prevention of transport
 CC disorders, neurological disorders, muscle disorders, immunological
 CC disorders and cell proliferative disorders, as well as in the creation of
 CC transgenic animals and during somatic or germline gene therapy to treat
 CC the disorders. The current sequence is that of a human TRICH protein of
 CC the invention.

XX Sequence 819 AA;

Query Match 82.9%; Score 29; DB 8; Length 819;

Best Local Similarity 62.5%; Pred. No. 1.8e+03;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|::|||
455 LLSVTIPM 462

Db

RESULT 37
AAB42511
ID AAB42511 standard; protein; 908 AA.

XX AC AAB42511;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF2275 polypeptide sequence SEQ ID NO:4550.

XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipariatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disease; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

XX OS Homo sapiens.

XX KW WO200058473-A2.

XX PD 05-OCT-2000.

XX PF 31-MAR-2000; 2000WO-US008621.

XX PR 31-MAR-1999; 99US-0127607P.

XX PR 02-APR-1999; 99US-0127636P.

XX PR 05-APR-1999; 99US-0127728P.

XX PR 30-MAR-2000; 2000US-00540763.

XX PA (CURA-) CURAGEN CORP.

XX PI Shinkets RA, Leach M;

XX DR WPI; 2000-602362/57.

XX DR N-PSDB; AAC76720.

XX PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.

XX PS Claim 11; Page 3747-3749; 5507pp; English.

XX CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipariatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester

CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX SQ Sequence 908 AA;

Query Match 82.9%; Score 29; DB 3; Length 908;
Best Local Similarity 62.5%; Pred. No. 2.1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|::|||
209 LLSVTIPM 216

Db

RESULT 38
AAO31015
ID AAO31015 standard; protein; 1307 AA.

XX AC AAO31015;

XX DT 06-OCT-2003 (first entry)

XX DE Human transporter and ion channel (TRICH)-23 protein.

XX KW Human; transporter and ion channel; TRICH; myotonic dystrophy; diabetes;
KW neurodegenerative disorder; Parkinson's disease; immunological disorder;
KW gastrointestinal disorder; Alzheimer's disease; hepatic disease; cancer;
KW cardiovascular disorder; atherosclerosis; endocrine disorder; catatonia;
KW systemic lupus erythematosus; Goodpasture's syndrome; muscular disorder;
KW Crohn's disease; transgenic animal; Grave's disease; infection; allergy;
KW leukaemia; scleroderma; gene therapy; transgenic; renal disorder;
KW cirrhosis.

XX OS Homo sapiens.

XX KW WO2003052119-A2.

XX PD 26-JUN-2003.

XX PF 10-DEC-2002; 2002WO-US039565.

XX PR 14-DEC-2001; 2001US-0340741P.

XX PR 25-JAN-2002; 2002US-0351359P.

XX PR 22-FEB-2002; 2002US-0359506P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Baughn MR, Chang H, Chien D, Duggan BM, Emerling BM, Corvad AE;
XX Griffin JA, Hafalia AJA, Hawkins PR, Jackson AA, Jiang X, Jin P;
XX Kable AE, Khare R, Lal PG, Lee EA, Lee S, Lee SY, Rankumar J;
XX Reddy R, Richardson TW, Ring HZ, Sprague WM, Swarnakar A, Tang YT;
XX Zebbarjadian Y;

XX WPI; 2003-533027/50.
XX DR N-PSDB; AAL62526.

XX PT New TRICH polypeptides, useful for diagnosing, preventing, and treating
XX disorders associated with abnormal expression or activity of TRICH, e.g.
XX neuromuscular, immunological, cardiovascular disorders, cancer and/or
XX infections.

XX PS Claim 1; Page 238-240; 273pp; English.

XX CC The invention provides human transporters and ion channels (TRICH)
XX polypeptides and polymucloides. Sequences of the invention are useful
XX for diagnosing, preventing and treating disorders associated with
XX abnormal expression or activity of TRICH such as neurodegenerative
XX disorders (e.g. Parkinson's disease, Alzheimer's disease), muscular
XX disorders (e.g. myotonic dystrophy, catatonia), endocrine disorders (e.g.
XX diabetes, Grave's disease), cancers (e.g. leukaemia, cervical or breast

CC cancers), immunological disorders (e.g. scleroderma, systemic lupus erythematosus, allergies), gastrointestinal disorders (e.g. Crohn's disease), renal disorders (e.g. Goodpasture's syndrome), infections (e.g. viral, bacterial, fungal, parasitic, protozoal, helminthic), CC cardiovascular disorders (e.g. atherosclerosis) or hepatic diseases (e.g. cirrhosis). TRICH polynucleotides can be used to create humanised animals or transgenic animals to model human disease. The invention is also used in gene therapy. The present sequence is human TRICH protein

XX Sequence 1307 AA;
SQ

Query Match 82.9%; Score 29; DB 6; Length 1307;
Best Local Similarity 62.5%; Pred. No. 3.1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 409 LLSVTIPM 416
:|:|:|

RESULT 39
AAE29913
ID AAE29913 standard; protein; 1353 AA.
XX AAE29913;
AC AAE29913;
XX

DT 24-FEB-2003 (first entry)
XX

DE Human transporter and ion channel (TRICH) protein #13.
XX

KW Human; transporter and ion channel; TRICH; neurodegenerative disorder;
KW Parkinson's disease; Alzheimer's disease; muscular disorder; transgenic;
KW myotonic dystrophy; cataract; endocrine disorder; diabetes; cytostatic;
KW Grave's disease; cancer; leukaemia; cervical; immunological; scleroderma;
KW systemic lupus erythematosus; allergy; gastrointestinal; Crohn's disease;
KW Goodpasture's syndrome; infection; cardiovascular; fungicide; nootropic;
KW hepatic disease; cirrhosis; gene therapy; uteropathic; anti-HIV; virucide;
KW atherosclerosis; antiparasitic; protozoacide; antibacterial.
XX

OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT Domain 130..158
FT Domain /note= "Transmembrane domain"
FT Domain 394..422
FT Domain /note= "Transmembrane domain"
FT Domain 448..473
FT Domain /note= "Transmembrane domain"
FT Domain 996..1024
FT Domain /note= "Transmembrane domain"
FT Domain 1033..1045
FT Domain /note= "EF-hand calcium binding domain"
FT Domain 1055..1083
FT Domain /note= "Transmembrane domain"
FT Domain 1093..1113
FT Domain /note= "Transmembrane domain"
FT Domain 1117..1137
FT Domain /note= "Transmembrane domain"
FT Domain 1163..1191
FT Domain /note= "Transmembrane domain"
XX

PN WO200277237-A2.
XX

XX

PD 03-OCT-2002.
XX

XX

PF 08-FEB-2002; 2002WO-US003657.
XX

XX

XX 09-FEB-2001; 2001US-0267892P.
PR 23-FEB-2001; 2001US-0271168P.
PR 02-MAR-2001; 2001US-0272890P.
PR 16-MAR-2001; 2001US-0276860P.
PR 23-MAR-2001; 2001US-0278255P.
PR 30-MAR-2001; 2001US-0280538P.

PR 25-JAN-2002; 2002US-0351359P.
XX

PA (INCY-) INCYTE GENOMICS INC.
XX

PI Lee EA, Ding L, Baughn MR, Tribouley CM, Bruns CM, Elliott VS;
PI Walia NK, Forsythe IJ, Raumann BE, Burford N, Lal PG, Thornton M;
PI Gandhi AR, Arvizu C, Yao MG, Yue H, Xu Y, Hafalia AUA, Ison CH;
PI Chen H;
XX

XX WPI; 2003-018931/01.
DR N-PSDB; AAD47360.
XX

XX New TRICH polypeptides, useful for diagnosing, preventing, and treating
FT disorders associated with an abnormal expression or activity of TRICH,
FT e.g. neuromuscular, immunological, cardiovascular disorders, cancer and
FT infection.
XX

PS Claim 1; Page 172-176; 214pp; English.
XX

XX The invention relates to human transporters and ion channels (TRICH) and
CC their nucleic acids. The sequences of the invention are useful in
CC diagnosing, preventing, and treating disorders associated with an
CC abnormal expression or activity of TRICH, such as neurodegenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), muscular
CC disorders (e.g. myotonic dystrophy, cataract), endocrine disorders (e.g.
CC diabetes, Grave's disease), cancers (e.g. leukaemia, cervical or breast
CC cancers), immunological disorders (e.g. scleroderma, systemic lupus
CC erythematosus, allergies), gastrointestinal disorders (e.g. Crohn's
CC disease), renal disorders (e.g. Goodpasture's syndrome), infections (e.g.
CC viral, bacterial, fungal, parasitic, protozoal, helminthic),
CC cardiovascular disorders (e.g. atherosclerosis), or hepatic diseases
CC (e.g. cirrhosis). TRICH or its fragments may also be used in screening
CC for compounds that specifically bind to and modulate its activity. TRICH
CC DNA can be used to create humanised animals or transgenic animals to
CC model human disease. It is also used in gene therapy. The present
CC sequence is human TRICH protein
XX

SQ Sequence 1353 AA;
Query Match 82.9%; Score 29; DB 6; Length 1353;
Best Local Similarity 62.5%; Pred. No. 3.3e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 455 LLSVTIPM 462
:|:|:|

RESULT 40
ADS75981
ID ADS75981 standard; protein; 1359 AA.
XX

AC ADS75981;
XX

XX 16-DEC-2004 (first entry)
DT

DE Human ATP8B3 protein, a nematode Dsc-3 protein homologue.
XX

KW antiarteriosclerotic; antilipemic; cardiovascular; anorectic;
KW gene therapy; lipid; lipoprotein; nematode; mutation; clk-1 gene;
KW mutagenesis; phenotype; defecation cycle; germline development;
KW soma development; embryonic development; post-embryonic development;
KW lipid metabolism disorder; atherosclerosis; dyslipidaemia;
KW cardiovascular disease; heart disease; blood vessel disease;
KW cerebrovascular disorder; obesity; Dsc-4 gene;
KW Modulator of Lipids and Lipoproteins.
XX

OS Homo sapiens.
XX

XX WO2004081231-A1.
PN

XX 23-SEP-2004.
PD

XX

```

PF 15-MAR-2004; 2004WO-CA000385.
XX
PR 14-MAR-2003; 2003US-0454925P.
XX
PA (UYMC-) UNIV MCGILL.
XX
PI Hekimi S, Shibata Y, Branicky R;
XX
XX WPI; 2004-690609/67.
DR
XX
XX Isolating a gene that modulates the level of a lipid or lipoprotein
PT comprises subjecting nematodes comprising mutation in the clk-1 gene to
PT mutagenesis.
PT
XX Disclosure; SEQ ID NO 11; 230pp; English.
XX
XX The invention relates to a method of isolating a gene that modulates the
CC level of a lipid or lipoprotein in nematodes by: (a) subjecting nematodes
CC that comprise at least one mutation in the clk-1 gene to mutagenesis to
CC produce test nematodes; (b) identifying test nematodes that exhibit any
CC of the phenotypes stated above that is modified as compared to the
CC phenotype of nematodes not subjected to mutagenesis; and (c) cloning the
CC gene that was mutated in step (a) from a test nematode of step (b), where
CC the phenotype is: (i) length of defecation cycle; (ii) rate of germline
CC development relative to rate of soma development; (iii) rate of embryonic
CC development; or (iv) rate of post-embryonic development. The method is
CC useful for isolating a gene that modulates the level of a lipid or
CC lipoprotein in nematodes. The nucleic acid, polypeptide, composition, and
CC methods are useful for preventing or treating lipid metabolism disorder,
CC e.g. atherosclerosis, dyslipidaemia, cardiovascular disease, heart
CC disease, blood vessel disease, cerebrovascular disorder, or obesity. This
CC sequence corresponds to a human Type IV P-Type ATPase which is a nematode
CC Dec-3 protein homologue.
XX
XX SQ Sequence 1359 AA;

Query Match 82.9%; Score 29; DB 8; Length 1359;
Best Local Similarity 62.5%; Pred. No. 3.3e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 382 LLSVTIPM 389
:|:|:|

RESULT 41
ADP99172
ID ADP99172 standard; protein; 1381 AA.
XX
XX ADP99172;
AC
XX 26-AUG-2004 (first entry)
DT
XX Human transporter and ion channel (TRICH) protein - SEQ ID 37.
DE
XX transporter; ion channel; TRICH; neuroprotective; respiratory; muscular;
XX antiparkinsonian; anticonvulsant; nootropic; neuroprotective;
XX antiinflammatory; endocrine; anti-HIV; antianaemic; immunosuppressive;
XX nephrotropic; hepatotropic; antipsoriatic; transport disorder;
XX neurological; muscle; immunological; cell proliferative; transgenic;
XX gene therapy; human.
XX
XX Homo sapiens.
OS
XX WO2004048599-A2.
FN
XX 10-JUN-2004.
PD
XX
XX 20-NOV-2003; 2003WO-US037572.
XX
XX 22-NOV-2002; 2002US-0428508P.
PR
XX 08-DEC-2002; 2002US-0431635P.
PR
XX 13-DEC-2002; 2002US-0433329P.
PR

24-JAN-2003; 2003US-042433P.
27-FEB-2003; 2003US-0451298P.
(INCY-) INCYTE CORP.
Tran UK, Jin P, Elliott VS, Yue H, Chawla NK, Yao MG;
Raumann BE, Khan FA, Arvizu CS, Warren BA, Griffin JA, Ison CH;
Hafalia AJA, Lal PG, Lee SY, Kable AE, Khare R, Swarnakar A;
Becha SD, Marquis JP, Favero KD, Wang JT, Richardson TW, Bhatia UG;
Burrill JD, Lee S, Blake JJ, Ho A, Zheng W, Nakamura LL;
WPI; 2004-441181/41.
DR N-PSDB; ADP99221.
XX
XX New isolated human transporters and ion channels, TRICH 1-49, useful for
PT diagnosing, treating, and preventing Duchenne's muscular dystrophy,
PT polymyositis, psoriasis, Alzheimer's disease, and Cushing's syndrome.
XX
XX Claim 1; SEQ ID NO 37; 292pp; English.
XX
XX The invention relates to a novel isolated transporter and ion channel
CC (TRICH) polypeptide comprising a polypeptide sequence as given in
CC specification or its biologically active or immunogenic fragment. The
CC polypeptide of the invention demonstrates neuroprotective, respiratory,
CC muscular, antiparkinsonian, anticonvulsant, nootropic, neuroprotective,
CC antiinflammatory, endocrine, anti-HIV, antianaemic, immunosuppressive,
CC nephrotropic, hepatotropic and antipsoriatic activities. The polypeptide
CC may be useful in the diagnosis, treatment and prevention of transport
CC disorders, neurological disorders, muscle disorders, immunological
CC disorders, and cell proliferative disorders, as well as in the creation of
CC transgenic animals and during somatic or germline gene therapy to treat
CC the disorders. The current sequence is that of a human TRICH protein of
CC the invention.
XX
XX SQ Sequence 1381 AA;

Query Match 82.9%; Score 29; DB 8; Length 1381;
Best Local Similarity 62.5%; Pred. No. 3.3e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
Db 455 LLSVTIPM 462
:|:|:|

RESULT 42
ABP54724
ID ABP54724 standard; peptide; 8 AA.
XX
XX ABP54724;
AC
XX 30-DEC-2002 (first entry)
DT
XX Hypoxia inducible factor 1-alpha peptide.
DE
XX Hypoxia inducible factor 1; HIF-1; human; transcription factor;
XX vasotropic; cardiant; thrombolytic; gene therapy.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Modified-site 4
FT /label= OTHER
FT /note= "hydroxyproline"
FT
XX WO200274980-A2.
FN
XX 26-SEP-2002.
PD
XX
XX 20-MAR-2002; 2002WO-US008946.
XX
XX 20-MAR-2001; 2001US-0277425P.
PR
XX 20-MAR-2001; 2001US-0277431P.
PR

```



```

PR 20-MAR-2001; 2001US-0277440P.
PR 09-NOV-2001; 2001US-032334P.
PR 09-NOV-2001; 2001US-0332493P.
PR 09-NOV-2001; 2001US-0345200P.
PR 20-DEC-2001; 2001US-0342598P.
PR 20-DEC-2001; 2001US-0345131P.
PR 20-DEC-2001; 2001US-0345132P.
PR 19-MAR-2002; 2002US-00101816.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
PA
XX Kaelin WG, Ivan M;
XX
XX WPI; 2002-740863/80.
DR
XX
XX New HIF-alpha mutain polypeptide and encoding nucleic acid, useful for
PT reducing, preventing or treating hypoxia or ischemia-related tissue
PT damage, and/or increasing angiogenesis and vascularization.
XX
XX
XX Example 3; Page 54; 108pp; English.
XX
XX The present sequence is that of a peptide fragment of hypoxia inducible
CC factor 1 (HIF 1) alpha subunit comprising amino acids 561-568 of the full
CC -length sequence. This region of HIF is highly conserved and, when
CC mutated to 8 consecutive alanines, leads to HIF stabilisation. An alanine
CC scan of this region showed that Leu-562 and Pro-564 are essential for
CC specific binding to Von Hippel-Landau (VHL) polypeptide. The invention
CC provides HIF mutants (see APP4716-20) which show decreased binding to
CC VHL and increased resistance to degradation in the presence of oxygen. A
CC claimed method of reducing, preventing or treating hypoxia or ischaemia-
CC related tissue damage comprises administering a specified HIF mutain,
CC including L562A and P564A mutants, or a nucleic acid encoding it. A
CC claimed method of increasing angiogenesis or vascularisation in a subject
CC also involves administering the HIF mutain or a nucleic acid encoding it
XX
XX Sequence 8 AA;
SQ
Query Match 80.0%; Score 28; DB 5; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 1 MLAPYIPM 8

RESULT 45
AAE30025
ID AAE30025 standard; peptide; 8 AA.
XX
XX AAE30025;
XX
XX 24-FEB-2003 (first entry)
DT
XX
XX Human hypoxia-inducible factor 1 (HIF-1) alpha derived peptide #1.
DE
XX
XX Entity localisation; light-generating fusion protein; LGP; diabetes;
KW hypoxic tissue imaging; detection; cancer; hypoxia; heart disease;
KW stroke; human; hypoxia-inducible factor 1; HIF-1 alpha.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 5
FT /note= "Hydroxylated"
FT
XX
XX WO200275278-A2.
XX
XX 26-SEP-2002.
XX
XX 20-MAR-2002; 2002WO-US008864.
XX
XX 20-MAR-2001; 2001US-0277425P.
XX

```

```

PR 20-MAR-2001; 2001US-0277431P.
PR 20-MAR-2001; 2001US-0277440P.
PR 09-NOV-2001; 2001US-032334P.
PR 09-NOV-2001; 2001US-0332493P.
PR 09-NOV-2001; 2001US-0345200P.
PR 20-DEC-2001; 2001US-0342598P.
PR 20-DEC-2001; 2001US-0345131P.
PR 20-DEC-2001; 2001US-0345132P.
PR 19-MAR-2002; 2002US-00101662.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
PA
XX Kaelin WG, Livingston DM, Kim W;
XX
XX WPI; 2003-018815/01.
DR
XX
XX Detecting localization of an entity e.g. hypoxic tissue, tumor, wound in
PT a subject, by using a light-generating fusion protein having a ligand
PT binding site and light-generating polypeptide moiety.
XX
XX Example 1; Fig 4B; 129pp; English.
XX
XX The invention relates to a method of detecting localisation of an entity
CC in a subject. The method involves administering to the subject a light-
CC generating fusion protein (LGP), or a cell expressing LGP, where LGP
CC comprises a ligand binding site and a light-generating polypeptide
CC moiety, and light generation of LGP changes upon binding of a ligand at
CC the ligand binding site, allowing for co-localisation of LGP and an
CC entity, and imaging localised LGP. The method is useful for detecting the
CC localisation of an entity, such as a molecule, macromolecule, polymer,
CC protein, antibody, protein complex, polysaccharide, nucleic acid,
CC particle, inert material, organelle, cell, embryo, microorganism,
CC bacteria, virus, fungus, prion, tumour, tissue, cellular environment
CC comprising damaged tissue, diseased tissue or hypoxic tissue, wound,
CC organ, proliferating cell and pathogen in a subject. It is particularly
CC useful for measuring prolyl hydroxylase activity; for imaging hypoxic
CC tissue and for detecting cancerous tissue. LGPs are useful for screening
CC modulators of activity or latency of (or predisposition to) disorders
CC such as hypoxia, cancer, diabetes, heart disease or stroke. The present
CC sequence is human hypoxia-inducible factor 1 (HIF-1) alpha derived
CC peptide. This sequence is used to illustrate the method of the invention
XX
XX Sequence 8 AA;
SQ
Query Match 80.0%; Score 28; DB 6; Length 8;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 1 MLAPYIPM 8

RESULT 46
AAE30028
ID AAE30028 standard; peptide; 8 AA.
XX
XX AAE30028;
XX
XX 24-FEB-2003 (first entry)
DT
XX
XX Human hypoxia-inducible factor 1 (HIF-1) alpha derived peptide #4.
DE
XX
XX Entity localisation; light-generating fusion protein; LGP; diabetes;
KW hypoxic tissue imaging; detection; cancer; hypoxia; heart disease;
KW stroke; human; hypoxia-inducible factor 1; HIF-1 alpha.
XX
XX Homo sapiens.
OS
XX
XX WO200275278-A2.
XX
XX 26-SEP-2002.
XX
XX

```

```
PF 20-MAR-2002; 2002WO-US008864.
XX
XX
PR 20-MAR-2001; 2001US-0277425P.
PR 20-MAR-2001; 2001US-0277431P.
PR 20-MAR-2001; 2001US-0277440P.
PR 09-NOV-2001; 2001US-0332334P.
PR 09-NOV-2001; 2001US-0332493P.
PR 09-NOV-2001; 2001US-0345200P.
PR 20-DEC-2001; 2001US-0342598P.
PR 20-DEC-2001; 2001US-0345131P.
PR 20-DEC-2001; 2001US-0345132P.
PR 19-MAR-2002; 2002US-00101662.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX PI Kaelin WG, Livingston DM, Kim W;
XX WPI; 2003-018815/01.
XX
XX Detecting localization of an entity e.g. hypoxic tissue, tumor, wound in
XX a subject, by using a light-generating fusion protein having a ligand
XX binding site and light-generating polypeptide moiety.
XX
XX Example 1; Fig 4B; 129pp; English.
XX
XX The invention relates to a method of detecting localisation of an entity
XX in a subject. The method involves administering to the subject a light-
XX generating fusion protein (LGP), or a cell expressing LGP, where LGP
XX comprises a ligand binding site and a light-generating polypeptide
XX moiety, and light generation of LGP changes upon binding of a ligand at
XX the ligand binding site, allowing for co-localisation of LGP and an
XX entity, and imaging localised LGP. The method is useful for detecting the
XX localisation of an entity, such as a molecule, macromolecule, polymer,
XX protein, antibody, protein complex, polysaccharide, nucleic acid,
XX particle, inert material, organelle, cell, embryo, microorganism,
XX bacteria, virus, fungus, prion, tumour, tissue, cellular environment
XX comprising damaged tissue, diseased tissue or hypoxic tissue, wound,
XX organ, proliferating cell and pathogen in a subject. It is particularly
XX useful for measuring prollyl hydroxylase activity; for imaging hypoxic
XX tissue and for detecting cancerous tissue. LGPs are useful for screening
XX modulators of activity or latency of (or predisposition to) disorders
XX such as hypoxia, cancer, diabetes, heart disease or stroke. The present
XX sequence is human hypoxia-inducible factor 1 (HIF-1) alpha derived
XX peptide. This sequence is used to illustrate the method of the invention
XX
XX Sequence 8 AA;
XX
XX Query Match 80.0%; Score 28; DB 6; Length 8;
XX Best Local Similarity 75.0%; Pred. No. 1.8e+06;
XX Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MLAXTIPM 8
XX |||: |||
XX Db 1 MLAPYIPM 8
XX
XX RESULT 47
XX AAO23489
XX ID AAO23489 standard; peptide; 12 AA.
XX
XX AC AAO23489;
XX
XX DT 12-FEB-2004 (first entry)
XX
XX DE Murine HIF-1alpha protein N-TAD region fragment (residues 560-571).
XX
XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
XX ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
XX antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
XX N-TAD.
XX
XX OS Mus sp.
XX
XX PN WO2003074560-A2.
XX
XX PD 12-SEP-2003.
XX
XX PF 05-MAR-2003; 2003WO-SE000372.
XX
XX PR 05-MAR-2002; 2002US-0361333P.
XX
XX PR 27-NOV-2002; 2002US-0429307P.
XX
XX PA (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
XX PI Pereira T, Poellinger L, Hellstroem M;
XX WPI; 2003-712876/67.
XX
XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
XX angiogenesis, or treating a condition associated with HIF-1alpha
XX underexpression in a cell, a group of cells, or an organism, e.g.
XX ischemia or inflammation.
XX
XX Claim 39; Fig 22; 96pp; English.
XX
XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
XX that has (a) an altered transactivation capacity and improved stability
XX at normoxia. The Hif-1alpha protein, polynucleotide, vector, and
XX pharmaceutical composition are useful for increasing angiogenesis,
XX interfering with a normal response to reoxygenation following hypoxia, or
XX treating a condition associated with HIF-1alpha underexpression in a
XX cell, a group of cells, or an organism, e.g. ischaemia, diabetic
XX retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
XX stroke. The proteins and pharmaceutical compositions are also useful for
XX mimicking the hypoxic response or artificially inducing a hypoxic
XX response in a cell, group of cells, or organism, inducing vascular
XX formation or vascular development in a cell or a group of cells.
XX increasing angiogenic activity in a cell, or influencing erythropoietin
XX production, metabolism, or an inflammatory response in a cell, a group of
XX cells, or an organism. Sequences AAO23482-489 represent fragments within
XX the N-TAD region of a murine HIF-1 alpha protein, that can be used for
XX treatment for hypoxic-related conditions
XX
XX Sequence 12 AA;
XX
XX Query Match 80.0%; Score 28; DB 7; Length 12;
XX Best Local Similarity 75.0%; Pred. No. 25;
XX Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MLAXTIPM 8
XX |||: |||
XX Db 1 MLAPYIPM 8
XX
XX RESULT 48
XX AAO23517
XX ID AAO23517 standard; peptide; 12 AA.
XX
XX AC AAO23517;
XX
XX DT 12-FEB-2004 (first entry)
XX
XX DE Murine HIF-1alpha protein N-TAD region fragment (residues 560-571).
XX
XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
XX ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
XX antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
XX N-TAD.
XX
XX OS Mus sp.
XX
XX PN WO2003074560-A2.
XX
XX PD 12-SEP-2003.
XX
XX PF 05-MAR-2003; 2003WO-SE000372.
```

```
XX 05-MAR-2002; 2002US-0361333P.
PR 27-NOV-2002; 2002US-0429307P.
XX (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
PI Pereira T, Poellinger L, Hellstrom M;
XX
XX WPI; 2003-712876/67.
XX
XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.
XX
XX Example 8; Fig 22; 96pp; English.
XX
XX The invention relates to a hypoxia-inducible factor (HIF)-1alpha protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1alpha protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1alpha underexpression in a
CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell or a group of cells,
CC increasing angiogenic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. Sequences AAO23510-517 represent fragments within
CC the N-TAD region of a murine HIF-1 alpha protein
XX
XX Sequence 12 AA;
SQ
Query Match 80.0%; Score 28; DB 7; Length 12;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAXTIPM 8
DB 1 MLAPYIPM 8
|||: |||
|||: |||
AAE30165
RESULT 49
AAE30165
ID AAE30165 standard; peptide; 14 AA.
XX
XX AAE30165;
XX
XX 24-FEB-2003 (first entry)
XX
XX Peptide #5 used to block HIF-1alpha/pVHL interaction.
XX
XX Hypoxia inducible factor; HIF hydroxylase; inflammatory disorder; cancer;
XX wound healing; ischaemia; transplantation; blood pressure; gene therapy.
XX
XX Unidentified.
XX
XX WO200274981-A2.
XX
XX 26-SEP-2002.
XX
XX 21-MAR-2002; 2002WO-GB001381.
XX
XX 21-MAR-2001; 2001GB-00007123.
XX
XX 02-AUG-2001; 2001GB-00018952.
XX
XX (ISIS-) ISIS INNOVATION LTD.
XX
XX Maxwell PH, Fugh CW, Ratcliffe PJ, Schofield CJ;
XX
```

```
DR WPI; 2003-018808/01.
XX
XX Novel isolated polypeptide useful for treating ischemia, wound healing,
PT auto-, allo-, and xeno-transplantation, systemic high blood pressure,
PT cancer, or inflammatory disorders.
XX
XX Example 1; Page 247; 256pp; English.
XX
XX The invention relates to polypeptides having hypoxia inducible factor
CC (HIF) hydroxylase activity, referred to as PHD polypeptides (PHD 1,2 and
CC 3) and nucleic acid molecules encoding such polypeptides. Polypeptides of
CC the invention are used for treating conditions such as ischaemia, wound
CC healing, auto-, allo-, and xeno-transplantation, systemic high blood
CC pressure, cancer, or inflammatory disorders. They are useful in anti-
CC sense regulation of the HIF hydroxylase activity and in particular HIF
CC prolyl hydroxylase activity within a cell. They are also used to identify
CC additional substrates of HIF hydroxylases. Sequences of the invention are
CC used to design double stranded RNAs for use in RNA interference. They are
CC used as therapeutic agents and in purification, isolation, or screening
CC methods involving immuno-precipitation techniques and for detecting
CC polypeptides in biological samples. The invention is useful in gene
CC therapy. The present sequence is a peptide used to block HIF-1alpha/pVHL
CC interaction. This sequence is used in the invention
XX
XX Sequence 14 AA;
SQ
Query Match 80.0%; Score 28; DB 6; Length 14;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAXTIPM 8
DB 6 MLAPYIPM 13
|||: |||
|||: |||
AAO23514
RESULT 50
AAO23514
ID AAO23514 standard; peptide; 14 AA.
XX
XX AAO23514;
XX
XX 12-FEB-2004 (first entry)
XX
XX Murine HIF-1alpha protein N-TAD region fragment (residues 560-573).
XX
XX HIF-1alpha; hypoxia-inducible factor -1alpha; vasotropic; antidiabetic;
XX ophthalmological; antiinflammatory; cardiant; antirheumatic; mouse;
XX antiarthritic; cerebroprotective; angiogenesis stimulator; gene therapy;
XX N-TAD.
XX
XX Mus sp.
XX
XX WO2003074560-A2.
XX
XX 12-SEP-2003.
XX
XX 05-MAR-2003; 2003WO-SE000372.
XX
XX 05-MAR-2002; 2002US-0361333P.
XX
XX 27-NOV-2002; 2002US-0429307P.
XX
XX (ANGI-) ANGIOGENETICS SWEDEN AB.
XX
XX Pereira T, Poellinger L, Hellstrom M;
XX
XX WPI; 2003-712876/67.
XX
XX New hypoxia-inducible factor (HIF)-1alpha protein, useful for increasing
PT angiogenesis, or treating a condition associated with HIF-1alpha
PT underexpression in a cell, a group of cells, or an organism, e.g.
PT ischemia or inflammation.
XX
XX Example 8; Fig 22; 96pp; English.
XX
```


XX The invention relates to a hypoxia-inducible factor (HIF)-1 α protein
CC that has (a) an altered transactivation capacity and improved stability
CC at normoxia. The HIF-1 α protein, polynucleotide, vector, and
CC pharmaceutical composition are useful for increasing angiogenesis,
CC interfering with a normal response to reoxygenation following hypoxia, or
CC treating a condition associated with HIF-1 α underexpression in a
CC cell, a group of cells, or an organism, e.g. ischaemia, diabetic
CC retinopathy, inflammation, coronary heart disease, rheumatoid arthritis,
CC stroke. The proteins and pharmaceutical compositions are also useful for
CC mimicking the hypoxic response or artificially inducing a hypoxic
CC response in a cell, group of cells, or organism, inducing vascular
CC formation or vascular development in a cell or a group of cells,
CC increasing angiogenic activity in a cell, or influencing erythropoietin
CC production, metabolism, or an inflammatory response in a cell, a group of
CC cells, or an organism. Sequences AAO23510-517 represent fragments within
CC the N-TAD region of a murine HIF-1 α protein
XX
SQ Sequence 14 AA;

Query Match 80.0%; Score 28; DB 7; Length 14;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLAXTIPM 8
|||: |||
Db 1 MLAPVIPM 8

Search completed: February 8, 2005, 20:15:07
Job time : 49.807 secs

This Page Blank (uspto)

Wed Feb 9 06:58:00 2005

PRIOR APPLICATION NUMBER: PCT/GB00/01826
PRIOR FILING DATE: 2000-05-12
PRIOR APPLICATION NUMBER: GB9911047.0
PRIOR FILING DATE: 1999-05-12
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Motif
US-10-901-583-8

Query Match 100.0%; Score 101; DB 16; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
Db 1 DLDLEMLAPYIPMDDDFQL 19

RESULT 2

US-10-287-670-25
Sequence 25, Application US/10287670
Publication No. US20030150005A1
GENERAL INFORMATION:
APPLICANT: Kaelin Jr., et al.
TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and

TITLE OF INVENTION: Diagnostic and
TITLE OF INVENTION: Therapeutic Methods Thereof
FILE REFERENCE: 20363-009CIP1
CURRENT APPLICATION NUMBER: US/10/287,670
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 10/101,662
PRIOR FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: 10/101,812
PRIOR FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: 10/101,816
PRIOR FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: 60/277,425
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/277,431
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/277,440
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/332,493
PRIOR FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: 60/345,131
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 60/342,598
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 60/345,132
PRIOR FILING DATE: 2001-12-20
Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 25
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-10-287-670-25

Query Match 100.0%; Score 101; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
Db 1 DLDLEMLAPYIPMDDDFQL 19

RESULT 3

US-10-901-583-9
Sequence 9, Application US/10901583
Publication No. US20050003452A1
GENERAL INFORMATION:
APPLICANT: Ratcliffe, Peter John
APPLICANT: Maxwell, Patrick Henry
APPLICANT: Pugh, Christopher William
TITLE OF INVENTION: Interaction Between the VHL Tumour
TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods
TITLE OF INVENTION: Relating Thereto
FILE REFERENCE: 3547.1000-000
CURRENT APPLICATION NUMBER: US/10/901,583
CURRENT FILING DATE: 2004-07-29
PRIOR APPLICATION NUMBER: US/09/959,873
PRIOR FILING DATE: 2001-11-09
PRIOR APPLICATION NUMBER: PCT/GB00/01826
PRIOR FILING DATE: 2000-05-12
PRIOR APPLICATION NUMBER: GB9911047.0
PRIOR FILING DATE: 1999-05-12
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 9
LENGTH: 34
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic peptide
US-10-901-583-9

Query Match 100.0%; Score 101; DB 16; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
Db 8 DLDLEMLAPYIPMDDDFQL 26

RESULT 4

US-09-922-958-5
Sequence 5, Application US/09922958
Patent No. US20020048794A1
GENERAL INFORMATION:
APPLICANT: POELLINGER, Lorenz
APPLICANT: PEREIRA, Teresa
APPLICANT: RUAS, Jorge
TITLE OF INVENTION: MECHANISM OF CONDITIONAL REGULATION OF THE HYPOXIA-INDUCIBLE FACT.
FILE REFERENCE: 3743/49008
CURRENT APPLICATION NUMBER: US/09/922,958
CURRENT FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: US 60/223,480
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 54
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-958-5

Query Match 100.0%; Score 101; DB 9; Length 54;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
Db 25 DLDLEMLAPYIPMDDDFQL 43

RESULT 5
US-10-425-833-8

; Sequence 8, Application US/10425833
; Publication No. US20040018606A1
; GENERAL INFORMATION:
; APPLICANT: Bohl, Jean Michael
; TITLE OF INVENTION: Control of protein systemic delivery of hypoxia using a tet-HIF1-
; TITLE OF INVENTION: chimeric transactivator
; FILE REFERENCE: 235748US0
; CURRENT APPLICATION NUMBER: US/10/425,833
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/376,269
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 409
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
US-10-425-833-8

Query Match 100.0%; Score 101; DB 15; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.9e-07; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
| | | | | | | | | | | | | | | | | | | | |
Db 235 DLDLEMLAPYIPMDDDFQL 253

RESULT 6
US-10-425-833-9
; Sequence 9, Application US/10425833
; Publication No. US20040018606A1
; GENERAL INFORMATION:
; APPLICANT: Bohl, Delphine
; TITLE OF INVENTION: Control of protein systemic delivery of hypoxia using a tet-HIF1-
; TITLE OF INVENTION: chimeric transactivator
; FILE REFERENCE: 235748US0
; CURRENT APPLICATION NUMBER: US/10/425,833
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/376,269
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 466
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
US-10-425-833-9

Query Match 100.0%; Score 101; DB 15; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
| | | | | | | | | | | | | | | | | | | | |
Db 235 DLDLEMLAPYIPMDDDFQL 253

RESULT 7
US-10-425-833-6
; Sequence 6, Application US/10425833
; Publication No. US20040018606A1
; GENERAL INFORMATION:
; APPLICANT: Bohl, Delphine
; APPLICANT: Heard, Jean Michael
; TITLE OF INVENTION: Control of protein systemic delivery of hypoxia using a tet-HIF1-
; TITLE OF INVENTION: chimeric transactivator

; FILE REFERENCE: 235748US0
; CURRENT APPLICATION NUMBER: US/10/425,833
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/376,269
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 538
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
US-10-425-833-6

Query Match 100.0%; Score 101; DB 15; Length 538;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
| | | | | | | | | | | | | | | | | | | | |
Db 364 DLDLEMLAPYIPMDDDFQL 382

RESULT 8
US-10-264-049-2606
; Sequence 2606, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2606
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2606

Query Match 100.0%; Score 101; DB 15; Length 542;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
| | | | | | | | | | | | | | | | | | | | |
Db 272 DLDLEMLAPYIPMDDDFQL 290

RESULT 9
US-10-425-833-7
; Sequence 7, Application US/10425833
; Publication No. US20040018608A1
; GENERAL INFORMATION:
; APPLICANT: Bohl, Delphine
; APPLICANT: Heard, Jean Michael
; TITLE OF INVENTION: Control of protein systemic delivery of hypoxia using a tet-HIF1-
; TITLE OF INVENTION: chimeric transactivator
; FILE REFERENCE: 235748US0
; CURRENT APPLICATION NUMBER: US/10/425,833
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/376,269
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 595

Query Match	100.0%;	Score 101;	DB 14;	Length 823;
Best Local Similarity	100.0%;	Pred. No. 4e-07;		
Best Overall Similarity	100.0%;	Mismatches 0;	Indels 0;	Gaps 0;
Conservative				

QY 1 DLDLEMLAPYIPMDDDFQL 19
|||
556 DLDLEMLAPYIPMDDDFQL 574
DB

RESULT 12
US-09-922-958-4
; Sequence 4, Application US/099222958
Patent No. US20020048794A1

APPLICANT: PEREIRA, Teresa
APPLICANT: RUEAS, Jorge
TITLE OF INVENTION: MECHANISM OF CONDITIONAL REGULATION OF THE HYPOXIA-INDUCIBLE FACTOR-1
TITLE OF INVENTION: THE VON HIPPEL-LINDAU TUMOR SUPPRESSOR PROTEIN
FILE REFERENCE: 3743/49008
CURRENT APPLICATION NUMBER: US/09/922.958
CURRENT FILING DATE: 2001-08-07

APPLICANT: wang, longcong
APPLICANT: Secrist, Heather
APPLICANT: Mohamath, Raodoh
APPLICANT: Indirias, Carol Y.
APPLICANT: Fan, Liqueun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER
STATE OF INVENTION: INVENTION

Query Match	100.0%;	Score 101;	DB 9;	Length 826;
Best Local Similarity	100.0%;	Pred. No. 4e-07;		
Matches 19:	Conservative 0;	Mismatches 0;	Indels	

QY 1 DLDLEMLAPYIPMDDDFQL 19
Dp 556 DLDLEMLAPYIPMDDDFQL 574

556 DLDLEMLAPYIPMDDDFQL 574

RESULT 14
US-09-736-457-330
; Sequence 330, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-330

Query Match 100.0%; Score 101; DB 9; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
|||||

Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 15
US-09-902-941-330
; Sequence 330, Application US/09902941
; Patent No. US2002017952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Matnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-330

Query Match 100.0%; Score 101; DB 9; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19

Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 16
US-09-849-626-330
; Sequence 330, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-330

Query Match 100.0%; Score 101; DB 9; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
|||||

Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 17
US-09-967-388-4
; Sequence 4, Application US/09967388
; Publication No. US20030103956A1
; GENERAL INFORMATION:
; APPLICANT: JEFFEREY M. ARBEIT
; TITLE OF INVENTION: USE OF HIF-1ALPHA VARIANTS TO ACCELERATE
; FILE REFERENCE: UCO77.001A
; CURRENT APPLICATION NUMBER: US/09/967,388
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HUMAN
US-09-967-388-4

Query Match 100.0%; Score 101; DB 10; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
|||||

Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 18
US-09-476-300-330
; Sequence 330, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-476-300-330

Query Match 100.0%; Score 101; DB 10; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDLEMLAPYIPMDDDFQL 19
|||||
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 19

US-10-028-158-23
; Sequence 23, Application US/10028158
; Publication No. US20020110833A1
; GENERAL INFORMATION:
; APPLICANT: Caniggia, Isabella
; APPLICANT: Post, Martin
; APPLICANT: Lye, Stephen
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF
; TITLE OF INVENTION: TROPHOBLAST
; FILE REFERENCE: 11757.38USWO
; CURRENT APPLICATION NUMBER: US/10/028,158
; CURRENT FILING DATE: 2001-12-20
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US/09/380,662
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-158-23

Query Match 100.0%; Score 101; DB 13; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDLEMLAPYIPMDDDFQL 19
|||||
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 20

US-10-101-812-10
; Sequence 10, Application US/10101812
; Publication No. US20020192737A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Ivan, Mircea
; TITLE OF INVENTION: Pharmaceuticals and Methods for Treating Hypoxia and
; TITLE OF INVENTION: Screening Methods Thereof
; FILE REFERENCE: 20363-014
; CURRENT APPLICATION NUMBER: US/10/101,812
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431

; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Consensus
; OTHER INFORMATION: Target Peptide
US-10-101-812-10

Query Match 100.0%; Score 101; DB 13; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDLEMLAPYIPMDDDFQL 19
|||||
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 21

US-10-101-662A-9
; Sequence 9, Application US/10101662A
; Publication No. US20030022198A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Livingston, David A
; APPLICANT: Kim, William
; TITLE OF INVENTION: Light Emitting Fusion Proteins and Diagnostic and
; TITLE OF INVENTION: Therapeutic Methods Thereof
; FILE REFERENCE: 20363-009
; CURRENT APPLICATION NUMBER: US/10/101,662A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-101-662A-9

Query Match 100.0%; Score 101; DB 14; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFOL 19
|||||
Db 556 DLDLEMLAPYIPMDDDFOL 574

RESULT 22

US-10-017-754-330
; Sequence 330, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-754-330

Query Match 100.0%; Score 101; DB 14; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFOL 19
|||||
Db 556 DLDLEMLAPYIPMDDDFOL 574

RESULT 23

US-10-115-987B-14
; Sequence 14, Application US/10115987B
; Publication No. US20030148521A1
; GENERAL INFORMATION:
; APPLICANT: Bell, John C.; Stojdl, David P.;
; APPLICANT: Gray, Douglas A.; Sonnenberg,
; APPLICANT: Nahum, Lichty, Brian
; TITLE OF INVENTION: Conditionally Replicative and
; TITLE OF INVENTION: Conditionally Active Viruses
; FILE REFERENCE: 42630-0001
; CURRENT APPLICATION NUMBER: US/10/115,987B
; CURRENT FILING DATE: 2002-03-03
; PRIOR APPLICATION NUMBER: US60/281,781
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: EditPad
; SEQ ID NO 14
; LENGTH: 826
; TYPE: PRT
; ORGANISM: homo Sapiens
; PUBLICATION INFORMATION:
; AUTHORS: Wang et al.
; JOURNAL: Proceedings of the National Academy of Sciences
; VOLUME: 92

; PAGES: 5510-5514
; DATE: 1995
US-10-115-987B-14

Query Match 100.0%; Score 101; DB 14; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFOL 19
|||||
Db 556 DLDLEMLAPYIPMDDDFOL 574

RESULT 24

US-10-287-670-9
; Sequence 9, Application US/10287670
; Publication No. US20030150005A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., et al.
; TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and
; TITLE OF INVENTION: Diagnostic and
; FILE REFERENCE: 20363-009C1P1
; CURRENT APPLICATION NUMBER: US/10/287,670
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 10/101,662
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,812
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,816
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-670-9

Query Match 100.0%; Score 101; DB 14; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFOL 19
|||||
Db 556 DLDLEMLAPYIPMDDDFOL 574

RESULT 25

US-10-113-872-330
; Sequence 330, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.

Wed Feb 9 06:58:00 2005

```
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-113-872-330

Query Match      100.0%; Score 101; DB 14; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
   |||||
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 26
US-10-423-419-2
; Sequence 2, Application US/10423419
; Publication No. US20030176349A1
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: STABLE HYPOXIA INDUCIBLE FACTOR-1 alpha
; TITLE OF INVENTION: AND METHOD OF USE
; FILE REFERENCE: JHUI500-1
; CURRENT APPLICATION NUMBER: US/10/423,419
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US/09/383,581
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/148,547
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-423-419-2

Query Match      100.0%; Score 101; DB 14; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
   |||||
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 27
US-10-283-017-330
; Sequence 330, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margaita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C20
; CURRENT APPLICATION NUMBER: US/10/283,017
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-283-017-330

Query Match      100.0%; Score 101; DB 15; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
   |||||
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 28
US-10-901-583-18
; Sequence 18, Application US/10901583
; Publication No. US20050003452A1
; GENERAL INFORMATION:
; APPLICANT: Ratcliffe, Peter John
; APPLICANT: Maxwell, Patrick Henry
; APPLICANT: Pugh, Christopher William
; TITLE OF INVENTION: Interaction Between the VHL Tumour
; TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods
; FILE REFERENCE: 3547.1000-000
; CURRENT APPLICATION NUMBER: US/10/901,583
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: US/09/959,873
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: PCT/GB00/01826
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: GB9911047.0
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-901-583-18

Query Match      100.0%; Score 101; DB 16; Length 826;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
   |||||
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 29
US-09-919-039-149
; Sequence 149, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
```

```
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1250434CD1
US-09-919-039-149

Query Match          100.0%; Score 101; DB 10; Length 827;
Best Local Similarity 100.0%; Pred. No. 4e-07; 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
    |||||
Db 557 DLDLEMLAPYIPMDDDFQL 575

RESULT 30
US-10-247-671-137
; Sequence 137, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 137
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1250434CD1
US-10-247-671-137

Query Match          100.0%; Score 101; DB 14; Length 827;
Best Local Similarity 100.0%; Pred. No. 4e-07; 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
    |||||
Db 557 DLDLEMLAPYIPMDDDFQL 575

RESULT 31
US-10-101-816-5
; Sequence 5, Application US/10101816
; Publication No. US20030045686A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Ivan, Mircea
; TITLE OF INVENTION: Muteins of Hypoxia Inducible Factor Alpha and Methods
; FILE REFERENCE: 20363-008
; CURRENT APPLICATION NUMBER: US/10/101,816
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIF Mutein
US-10-101-816-2

Query Match          92.1%; Score 93; DB 14; Length 826;
Best Local Similarity 94.7%; Pred. No. 7.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIF Mutein
US-10-101-816-5

Query Match          95.0%; Score 96; DB 14; Length 826;
Best Local Similarity 94.7%; Pred. No. 2.4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
    |||||
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 32
US-10-101-816-2
; Sequence 2, Application US/10101816
; Publication No. US20030045686A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Ivan, Mircea
; TITLE OF INVENTION: Muteins of Hypoxia Inducible Factor Alpha and Methods
; FILE REFERENCE: 20363-008
; CURRENT APPLICATION NUMBER: US/10/101,816
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/332,334
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIF Mutein
US-10-101-816-2

Query Match          92.1%; Score 93; DB 14; Length 826;
Best Local Similarity 94.7%; Pred. No. 7.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 DLDLEMLAPYIPMDDDFQL 19
Db 556 DLDLEMLAAYIPMDDDFQL 574

RESULT 33

US-10-101-662A-15
; Sequence 15, Application US/10101662A
; Publication No. US2003002198A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Livingston, David A
; APPLICANT: Kim, William
; TITLE OF INVENTION: Light Emitting Fusion Proteins and Diagnostic and
; TITLE OF INVENTION: Therapeutic Methods Thereof
; FILE REFERENCE: 20363-009
; CURRENT APPLICATION NUMBER: US/101/101,662A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 15

LENGTH: 20

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: VARIANT

LOCATION: (9)

OTHER INFORMATION: Wherein Xaa is hydroxyproline

US-10-287-670-15

Query Match 91.1%; Score 92; DB 14; Length 20;
Best Local Similarity 94.7%; Pred. No. 1.7e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19

Db 1 DLDLEMLAAYIPMDDDFQL 19

RESULT 35

US-10-313-643A-5
; Sequence 5, Application US/10313643A
; Publication No. US2003015303A1
; GENERAL INFORMATION:
; APPLICANT: Klaus, Stephen J.
; APPLICANT: Lin, Al Y.
; APPLICANT: Neff, Thomas B.
; APPLICANT: Wang, Qingjian
; APPLICANT: Arend, Michael P.
; APPLICANT: Flippin, Lee A.
; APPLICANT: Melekhov, Alexey G.
; TITLE OF INVENTION: METHODS OF INCREASING ENDOGENOUS ERYTHROPOIETIN (EPO)
; FILE REFERENCE: FP0601 US
; CURRENT APPLICATION NUMBER: US/10/313,643A
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/349,659
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/386,488
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/337,082
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US 60/359,683
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5

LENGTH: 19

TYPE: PRT

ORGANISM: artificial sequence

FEATURE:

OTHER INFORMATION: synthetic peptide

US-10-313-643A-5

Query Match 91.1%; Score 92; DB 14; Length 20;
Best Local Similarity 94.7%; Pred. No. 1.7e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19

Db 1 DLDLEMLAAYIPMDDDFQL 19

RESULT 34

US-10-287-670-15
; Sequence 15, Application US/10287670
; Publication No. US20030150005A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., et al.
; TITLE OF INVENTION: Transgenic Animals Expressing Light Emitting Fusion Proteins and
; TITLE OF INVENTION: Diagnostic and
; TITLE OF INVENTION: Therapeutic Methods Thereof
; FILE REFERENCE: 20363-009CIP1
; CURRENT APPLICATION NUMBER: US/10/287,670
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 10/101,662
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 10/101,812

```
Query Match      88.1%; Score 89; DB 14; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.7e-07;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
    ||||| ||||| ||||| |||||
Db 1 DLDLEALAPYIPADDDFQL 19

RESULT 36
US-10-313-551A-5
; Sequence 5, Application US/10313551A
; Publication No. US20030176317A1
; GENERAL INFORMATION:
; APPLICANT: Guenzler-Pukall, Volkmar
; APPLICANT: Neff, Thomas B.
; APPLICANT: Wang, Qingjian
; APPLICANT: Arend, Michael
; APPLICANT: Flippin, Lee A.
; APPLICANT: Melikhov, Alexey G.
; TITLE OF INVENTION: STABILIZATION OF HYPOXIA INDUCIBLE FACTOR (HIF) ALPHA
; FILE REFERENCE: FP0600 US
; CURRENT APPLICATION NUMBER: US/10/313,551A
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/337,082
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US 60/359,683
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: US 60/349,659
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/386,488
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-313-551A-5

Query Match      88.1%; Score 89; DB 14; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.7e-07;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
    ||||| ||||| ||||| |||||
Db 1 DLDLEALAPYIPADDDFQL 19

RESULT 37
US-10-101-816-6
; Sequence 6, Application US/10101816
; Publication No. US20030045686A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Ivan, Mircea
; TITLE OF INVENTION: Mutins of Hypoxia Inducible Factor Alpha and Methods
; FILE REFERENCE: 20363-008
; CURRENT APPLICATION NUMBER: US/10/101,816
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIF Mutain
US-10-101-816-7

Query Match      87.1%; Score 88; DB 14; Length 826;
Best Local Similarity 89.5%; Pred. No. 4.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
    ||||| ||||| ||||| |||||
Db 1 DLDLEALAPYIPADDDFQL 19
```

```
Query Match      87.1%; Score 88; DB 14; Length 826;
Best Local Similarity 89.5%; Pred. No. 4.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
    ||||| ||||| ||||| |||||
Db 1 DLDLEALAPYIPADDDFQL 574

RESULT 38
US-10-101-816-7
; Sequence 7, Application US/10101816
; Publication No. US20030045686A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Ivan, Mircea
; TITLE OF INVENTION: Mutins of Hypoxia Inducible Factor Alpha and Methods
; FILE REFERENCE: 20363-008
; CURRENT APPLICATION NUMBER: US/10/101,816
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIF Mutain
US-10-101-816-7

Query Match      87.1%; Score 88; DB 14; Length 826;
Best Local Similarity 89.5%; Pred. No. 4.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMDDDFQL 19
    ||||| ||||| ||||| |||||
Db 1 DLDLEALAPYIPADDDFQL 19
```

```
Db      556 DLLEMAAAIPIPMDDDFOL 574

RESULT 39
US-10-296-115-933
; Sequence 933, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hysseq Inc
; TITLE OF INVENTION: NO. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296.115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 933
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-933

Query Match      86.1%; Score 87; DB 15; Length 297;
Best Local Similarity 94.4%; Pred. No. 2e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 LDLEMLAPIPMDDDFOL 19
      ||||| ||||| |||||
Db      166 LDLEMLAPIISMDDDFOL 183

RESULT 40
US-10-154-386-2
; Sequence 2, Application US/10154386
; Publication No. US20030026793A1
; GENERAL INFORMATION:
; APPLICANT: Angiogene Inc.
; APPLICANT: Guy, Louis-Georges
; TITLE OF INVENTION: HIPOXIA INDUCING FACTORS AND USBS THEREOF FOR INDUCING ANGIOGENESIS
; FILE REFERENCE: 5600-81
; CURRENT APPLICATION NUMBER: US/10/154.386
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/292,630
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/354529
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-386-2

Query Match      86.1%; Score 87; DB 14; Length 705;
Best Local Similarity 94.4%; Pred. No. 5.2e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 LDLEMLAPIPMDDDFOL 19
      ||||| ||||| |||||
Db      521 LDLEMLAPIISMDDDFOL 538

RESULT 41
US-10-121-235-2
; Sequence 2, Application US/10121235
; Publication No. US20030032609A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Maemura, Koji

; APPLICANT: Hsieh, Chung-Ming
; TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS
; FILE REFERENCE: 05433/037001
; CURRENT APPLICATION NUMBER: US/10/121,235
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 09/374,454
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: US 60/096,515
; PRIOR FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-235-2

Query Match      75.7%; Score 76.5; DB 14; Length 205;
Best Local Similarity 80.0%; Pred. No. 0.00058;
Matches 16; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy      1 LDLEMLAPIPMD-DDFOL 19
      :||| :||| :||| :|||
Db      38 ELDTLAPIPMDDGFOL 57

RESULT 42
US-10-121-235-6
; Sequence 6, Application US/10121235
; Publication No. US20030032609A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Maemura, Koji
; APPLICANT: Hsieh, Chung-Ming
; TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS
; FILE REFERENCE: 05433/037001
; CURRENT APPLICATION NUMBER: US/10/121,235
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 09/374,454
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: US 60/096,515
; PRIOR FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-121-235-6

Query Match      75.7%; Score 76.5; DB 14; Length 870;
Best Local Similarity 80.0%; Pred. No. 0.0029;
Matches 16; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy      1 LDLEMLAPIPMD-DDFOL 19
      :||| :||| :||| :|||
Db      523 ELDTLAPIPMDDGFOL 542

RESULT 43
US-10-101-816-4
; Sequence 4, Application US/10101816
; Publication No. US20030045686A1
; GENERAL INFORMATION:
; APPLICANT: Kaelin Jr., William G
; APPLICANT: Ivan, Mircea
; TITLE OF INVENTION: Mutins of Hypoxia Inducible Factor Alpha and Methods
; FILE REFERENCE: 20363-008
; CURRENT APPLICATION NUMBER: US/10/101,816
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/277,425
; PRIOR FILING DATE: 2001-03-20
```

; PRIOR APPLICATION NUMBER: 60/277,431
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/277,440
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/332,493
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,200
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,131
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/342,598
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/345,132
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/332,334
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIF Mutein
US-10-101-816-4

Query Match 60.9%; Score 61.5; DB 14; Length 870;
Best Local Similarity 70.0%; Pred. No. 0.63;
Matches 14; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 DLDLEMLAPYIPMD-DDFOL 19
:|||||:|||||:|
Db 523 ELDLETLAAYIPMDGEGFOL 542

RESULT 44
US-10-080-608A-37
; Sequence 37, Application US/10080608A
; Publication No. US20030198956A1
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hymen, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/10/080,608A
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 1507
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-080-608A-37

Query Match 53.5%; Score 54; DB 14; Length 1507;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 MLAPYIPMDDDFOL 19
:|||||:|||||:|
Db 847 MRAPYIPIDDDMPL 860

RESULT 45
US-10-370-685-126
; Sequence 126, Application US/10370685
; Publication No. US20030215903A1
; GENERAL INFORMATION:
; APPLICANT: Hymen, Paul
; APPLICANT: Goldberg, Edward
; TITLE OF INVENTION: Nanostructures Containing PNA Joining and Functional Elements
; FILE REFERENCE: NAN.P-004

; CURRENT APPLICATION NUMBER: US/10/370,685
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: 10/080,608
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 1507
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-370-685-126

Query Match 53.5%; Score 54; DB 15; Length 1507;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 MLAPYIPMDDDFOL 19
:|||||:|||||:|
Db 847 MRAPYIPIDDDMPL 860

RESULT 46
US-10-437-963-147229
; Sequence 147229, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147229
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_47777C.1.pep
US-10-437-963-147229

Query Match 44.6%; Score 45; DB 16; Length 162;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLDLEMLAPYIPMD 14
:|||||:|||||:|
Db 132 EMDIETLNPYPFGD 145

RESULT 47
US-10-437-963-147230
; Sequence 147230, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B

Wed Feb 9 06:58:00 2005

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 147230
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4778C.1.pep
US-10-437-963-147230

Query Match      44.6%; Score 45; DB 16; Length 172;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      1 DLDLEMLAPYIPMD 14
       :.:|:|:|:|:|:|
Db      142 EMDIETLNPYVPGD 155

RESULT 48
US-10-369-493-3670
; Sequence 3670, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3670
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(793)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3670

Query Match      44.6%; Score 45; DB 15; Length 793;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      6 MLAPYIPMDDD 16
       :.:|:|:|:|:|:|
Db      277 MLAGYLPFDDDD 287

RESULT 49
US-10-369-493-22401
; Sequence 22401, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3670
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(793)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-22401

Query Match      44.6%; Score 45; DB 15; Length 800;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      6 MLAPYIPMDDD 18
       :.:|:|:|:|:|:|
Db      245 MLAGYLPWDDDDHE 257

RESULT 50
US-10-618-581-6
; Sequence 6, Application US/10618581
; Publication No. US20040077524A1
; GENERAL INFORMATION:
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR SELECTIVELY INHIBITING FUNGAL GROWTH
; FILE REFERENCE: 034536/0123
; CURRENT APPLICATION NUMBER: US/10/618,581
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,624
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-618-581-6

Query Match      44.6%; Score 45; DB 15; Length 800;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      6 MLAPYIPMDDD 18
       :.:|:|:|:|:|:|
Db      245 MLAGYLPWDDDDHE 257

Search completed: February 9, 2005, 06:35:42
Job time : 256.667 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 20:05:10 ; Search time 8.2807 Seconds
(without alignments)
92.955 Million cell updates/sec

Title: US-10-032-361-4
Perfect score: 35
Sequence: 1 MLAXTIPM 8

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 65 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	91.4	321	F82073	sigma-E factor reg
2	31	88.6	686	A45483	chloride channel, C
3	31	88.6	687	D57713	chloride channel, C
4	31	88.6	687	C57713	chloride channel C
5	31	88.6	687	A57713	chloride channel C
6	30	85.7	650	C95101	PTS system, fructo
7	30	85.7	650	D97969	phosphotransferase
8	30	85.7	739	B82552	phage-related prot
9	30	85.7	739	F82769	phage-related tail
10	29	82.9	120	E70739	probable pks14 pro
11	29	82.9	227	E97526	30S ribosomal prot
12	29	82.9	243	C72776	hypothetical prote
13	29	82.9	255	A92745	30S ribosomal prot
14	29	82.9	261	T05447	hypothetical prote
15	29	82.9	405	A75459	sodium extrusion p
16	29	82.9	430	F96686	unknown protein F1
17	29	82.9	462	G87500	hypothetical prote
18	29	82.9	463	AH1383	amino acid transpo
19	29	82.9	463	AI1758	amino acid transpo
20	29	82.9	493	B44761	6-aminohexanoate-c
21	29	82.9	493	A44761	6-aminohexanoate-c
22	29	82.9	493	A44761	6-aminohexanoate-c
23	29	82.9	515	G75267	trk potassium upta
24	29	82.9	721	D87553	ABC transporter, p
25	29	82.9	729	D72079	protein C43E11.6
26	29	82.9	729	H86544	conserved hypotet
27	29	82.9	729	H86544	hypothetical prote
28	28	80.0	119	S72786	hypothetical prote
29	28	80.0	270	E86662	phosphonate ABC tr

30	28	80.0	407	2	D95882	probable efflux pr
31	28	80.0	413	2	A87305	Na+/H+ antiporter
32	28	80.0	465	2	S76464	hypothetical prote
33	28	80.0	504	2	S46752	hypothetical prote
34	28	80.0	524	2	H84141	L-lactate permease
35	28	80.0	534	2	T30629	hypothetical prote
36	28	80.0	545	2	A12327	hypothetical prote
37	28	80.0	698	2	H75577	probable drug tran
38	28	80.0	810	2	JC4837	hypoxia-inducible
39	28	80.0	811	2	JC7619	hypoxia-inducible
40	28	80.0	813	2	JC5809	hypoxia-inducible
41	28	80.0	826	2	I38972	probable membrane
42	28	80.0	908	2	S51293	hypothetical prote
43	27	77.1	89	2	G83514	conserved hypotet
44	27	77.1	109	2	I52333	G1 phase-specific
45	27	77.1	136	2	AB1127	hypothetical prote
46	27	77.1	136	2	AG1487	hypothetical prote
47	27	77.1	150	2	F95233	universal stress p
48	27	77.1	171	2	H98097	conserved hypotet
49	27	77.1	209	2	G70720	probable pgsA2 pro
50	27	77.1	220	2	S40931	hypothetical prote
51	27	77.1	240	2	E70771	hypothetical prote
52	27	77.1	252	2	T09387	NMS32/34 protein -
53	27	77.1	258	2	E91201	type III secretion
54	27	77.1	260	2	A86048	escT (imported) -
55	27	77.1	260	2	AH2158	hypothetical prote
56	27	77.1	269	2	D82060	conserved hypotet
57	27	77.1	270	2	T19033	hypothetical prote
58	27	77.1	300	2	B56118	vetispiradiene syn
59	27	77.1	311	2	A72487	hypothetical prote
60	27	77.1	317	2	D83963	methionyl-tRNA for
61	27	77.1	328	2	T11938	NADH2 dehydrogenas
62	27	77.1	334	2	T24310	hypothetical prote
63	27	77.1	350	2	C56118	vetispiradiene syn
64	27	77.1	378	2	C83190	lipid A-disacchari
65	27	77.1				

ALIGNMENTS

RESULT 1

F82073
sigma-E factor regulatory protein RseB VC2465 [imported] - Vibrio cholerae (strain N1696)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82073
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406: 477-483, 2000
A:Title: DNA-Sequence-of-both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20408833; PMID:10952301
A:Accession: F82073
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <HEI>
A:Cross-references: UNIPROT:Q9KPA8; GB:AE004316; GB:AE003852; NID:g9657034; PIDN:AAF9560
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2465
C:Map position: 1
C:Superfamily: conserved hypothetical protein HI0630

Query Match 91.4%; Score 32; DB 2; Length 321;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MLAXTIPM-8
Db 106 MLAXTIPM 113

hydrophobic

RESULT 2

A45483
Chloride channel, CLC-K1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Dec-2002
C:Accession: A45483
R:Uchida, S.; Sasaki, S.; Furukawa, T.; Hiraoka, M.; Imai, T.; Hirata, Y.; Marumo, F.
J. Biol. Chem. 268, 3821-3824, 1993
A:Title: Molecular cloning of a chloride channel that is regulated by dehydration and ex
A:Reference number: A45483; MUID:93179373; PMID:7680033
A:Accession: A45483
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-686 <UCH>
A:Experimental source: kidney
A:Note: sequence extracted from NCBI backbone (NCBIP:125743)
C:Superfamily: Chloride channel; CBS homology
C:Keywords: transmembrane protein
F:630-677/Domain: CBS homology <CBS>

Query Match 88.6%; Score 31; DB 2; Length 686;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|||||
Db 414 ILATTIPM 421

RESULT 3

D57713
Chloride channel CLC-Kb - human
C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: D57713
R:Kieferle, S.; Fong, P.; Bens, M.; Vandewalle, A.; Jentsch, T.J.
Proc. Natl. Acad. Sci. U.S.A. 91, 6943-6947, 1994
A:Title: Two highly homologous members of the CLC chloride channel family in both rat an
A:Reference number: A57713; MUID:94316614; PMID:8041726
A:Accession: D57713
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-687 <KIE>
A:Cross-references: UNIPROT:P51801; GB:Z30644; NID:9521073; PIDN:CAA83121.1; PID:9521074
C:Superfamily: Chloride channel; CBS homology
C:Keywords: chloride channel
F:631-678/Domain: CBS homology <CBS>

Query Match 88.6%; Score 31; DB 2; Length 687;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|||||
Db 414 ILATTIPM 421

RESULT 4

C57713
Chloride channel CLC-ka - human
C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: C57713
R:Kieferle, S.; Fong, P.; Bens, M.; Vandewalle, A.; Jentsch, T.J.
Proc. Natl. Acad. Sci. U.S.A. 91, 6943-6947, 1994
A:Title: Two highly homologous members of the CLC chloride channel family in both rat an
A:Reference number: A57713; MUID:94316614; PMID:8041726
A:Accession: C57713
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-687 <KIE>
A:Cross-references: UNIPROT:P51800; GB:Z30643; NID:9521071; PIDN:CAA83120.1; PID:9521072
C:Superfamily: Chloride channel; CBS homology

C:Keywords: chloride channel

Query Match 88.6%; Score 31; DB 2; Length 687;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|||||
Db 414 ILATTIPM 421

RESULT 5

A57713
Chloride channel CLC-K1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: A57713; S49333
R:Kieferle, S.; Fong, P.; Bens, M.; Vandewalle, A.; Jentsch, T.J.
Proc. Natl. Acad. Sci. U.S.A. 91, 6943-6947, 1994
A:Title: Two highly homologous members of the CLC chloride channel family in both rat an
A:Reference number: A57713; MUID:94316614; PMID:8041726
A:Accession: A57713
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-687 <KIE>
A:Cross-references: UNIPROT:Q06393; EMBL:Z34291; NID:9556670; PIDN:CAA84064.1; PID:95566
C:Superfamily: Chloride channel; CBS homology
C:Keywords: chloride channel
F:631-678/Domain: CBS homology <CBS>

Query Match 88.6%; Score 31; DB 2; Length 687;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|||||
Db 414 ILATTIPM 421

RESULT 6

C95101
PTS system, fructose specific IIAC components [imported] - Streptococcus pneumoniae (st
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: C95101
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95101
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-650 <KUR>
A:Cross-references: UNIPROT:Q97RES; GB:AB005672; PIDN:AAK75004.1; PID:gl4972350; GSPDB:G
C:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0877
C:Superfamily: phosphotransferase system enzyme II, fructose-specific; phosphotransferas.

Query Match 85.7%; Score 30; DB 2; Length 650;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|||||
Db 458 MLAVNIPM 465

RESULT 7

D97969

phosphotransferase system enzyme II (EC 2.7.1.69), fructose-specific fruA [imported] - S
 C:Species: Streptococcus pneumoniae
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C:Accession: D97969
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: D97969
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-650 <KUR>
 A:Cross-references: UNIPROT:Q8DQ95; GB:AE007317; PIDN:AAK99584.1; PID:gl5458376; GSPDB:G
 C:Genetics:
 A:Gene: fruA
 C:Superfamily: phosphotransferase system enzyme II, fructose-specific; phosphotransferase
 C:Keywords: phosphotransferase

Query Match 85.7%; Score 30; DB 2; Length 650;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
 |||: |||
 Db 458 MLAVNIPM 465

RESULT 8
 B82552
 phage-related protein XF2482 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: B82552
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: B82552
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-739 <SIM>
 A:Cross-references: UNIPROT:Q9PPE6; GB:AE003849; NID:g9105613; PIDN:AAF8354
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.Y.;
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0730

Query Match 85.7%; Score 30; DB 2; Length 739;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
 |||: |||
 Db 167 LARTIPM 173

RESULT 10
 E70739
 probable pks14 protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: E70739
 R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: E70739
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-120 <COL>
 A:Cross-references: UNIPROT:Q11012; GB:Z75555; GB:AL123456; NID:g3261608; PIDN:CAA99975.
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: pks14

Query Match 82.9%; Score 29; DB 2; Length 120;
 Best Local Similarity 62.5%; Pred. No. 33;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
 |||: |||
 Db 94 LLACTIPL 101

Query Match 85.7%; Score 30; DB 2; Length 739;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
 |||: |||
 Db 167 LARTIPM 173

RESULT 11
E97526
30S ribosomal protein S2 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: E97526
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: E97526
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-227 <KUR>
A:CROSS-references: GB:AE007869; PIDN:AAK87166.1; PID:g15156440; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_2539
A:Map position: circular chromosome
C:Superfamily: Escherichia coli ribosomal protein S2

Query Match 82.9%; Score 29; DB 2; Length 227;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
|||:|
Db 17 LAQTVP 23

RESULT 12
C72776
hypothetical protein APE0197 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: C72776
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Maesuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: C72776
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-243 <KAW>
A:CROSS-references: UNIPROT:Q9YFQ2; DBJ:AP000058; NID:g5103388; PIDN:BAA79109.1; PID:dl
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0197

Query Match 82.9%; Score 29; DB 2; Length 243;
Best Local Similarity 71.4%; Pred. No. 69;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
|||:|
Db 92 MLAGTVP 98

RESULT 13
AF2745
30S ribosomal protein S2 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AF2745
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AF2745
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <KUR>
A:CROSS-references: UNIPROT:Q8UFM3; GB:AE008688; PIDN:AAL42380.1; PID:g17739789; GSPDB:GI
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: rpsB
A:Map position: circular chromosome
C:Superfamily: Escherichia coli ribosomal protein S2

Query Match 82.9%; Score 29; DB 2; Length 255;
Best Local Similarity 71.4%; Pred. No. 72;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
|||:|
Db 45 LAQTVP 51

RESULT 14
T05447
hypothetical protein F7K2.110 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05447
R:Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuelle submitted to the Protein Sequence Database, November 1998
A:Reference number: Z15416
A:Accession: T05447
A:Molecule type: DNA
A:Residues: 1-261 <BEV>
A:CROSS-references: UNIPROT:Q9SUW6; EMBL:AL033545
A:Experimental source: cultivar Columbia; BAC clone F7K2
C:Genetics:
A:Map position: 4
A:Introns: 50/3
A:Note: F7K2.110
C:Superfamily: Arabidopsis thaliana hypothetical protein T20K12.110

Query Match 82.9%; Score 29; DB 2; Length 261;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
|||:|
Db 255 MLAGTVP 261

RESULT 15
A75459
sodium extrusion protein NatB - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: A75459
R:White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M. Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75459
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <WHI>
A:CROSS-references: UNIPROT:Q9RVU6; GB:AB001945; GB:AE00513; NID:g6458642; PIDN:AAF1050
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0926
A:Map position: 1

Query Match 82.9%; Score 29; DB 2; Length 405;

Best Local Similarity 62.5%; Pred. No. 1.2e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
|||: -|||
Db 183 MLAFVPM 190

RESULT 16
F96686
unknown protein F15E12.7 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F96686
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Chung, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96686
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-430 <STO>
A:Cross-references: UNIPROT:Q9C8C9; GB:AE005173; NID:g11038483; PIDN:AAG27762.1; GSPDB:G
C:Genetics:
A:Gene: F15E12.7
A:Map position: 1

Query Match 82.9%; Score 29; DB 2; Length 430;
Best Local Similarity 71.4%; Pred. No. 1.2e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
|||: -|||
Db 265 LAYTIPM 271

RESULT 17
G87500
hypothetical protein CC2028 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: G87500
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: G87500
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <STO>
A:Cross-references: UNIPROT:Q9A6Q6; GB:AE005673; NID:g13423503; PIDN:AAK24003.1; GSPDB:G
C:Genetics:
A:Gene: CC2028

Query Match 82.9%; Score 29; DB 2; Length 462;
Best Local Similarity 71.4%; Pred. No. 1.3e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
|||: -|||
Db 423 LAPTIPM 429

RESULT 18

AE1383
amino acid transporter homolog lmo2469 [imported] - Listeria monocytogenes (strain EGD-e
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1383
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative Genomics of Listeria species.
A:Reference number: AE1077; MUID:21537279; PMID:11679669
A:Accession: AE1383
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <GLA>
A:Cross-references: UNIPROT:Q8V4H3; GB:NC_003210; PIDN:CAD00547.1; PID:g16411957; GSPDB:G
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2469
C:Superfamily: arginine permease

Query Match 82.9%; Score 29; DB 2; Length 463;
Best Local Similarity 50.0%; Pred. No. 1.3e+02; Mismatches 4; Indels 0; Gaps 0;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
|||: -|||
Db 365 LIATVPM 372

RESULT 19
AH1758
amino acid transporter homolog lin2613 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1758
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative Genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1758
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <GLA>
A:Cross-references: UNIPROT:Q928C3; GB:AL592022; PIDN:CAC97840.1; PID:g16415150; GSPDB:G
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2613
C:Superfamily: arginine permease

Query Match 82.9%; Score 29; DB 2; Length 463;
Best Local Similarity 50.0%; Pred. No. 1.3e+02; Mismatches 4; Indels 0; Gaps 0;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
|||: -|||
Db 365 LIATVPM 372

RESULT 20
B44761
6-aminohexanoate-cyclic-dimer hydrolase (EC 3.5.2.12) precursor - Flavobacterium sp. (st
C:Species: Flavobacterium sp.
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Oct-1999
C:Accession: B44761
R:Tsuchiya, K.; Fukuyama, S.; Kanzaki, N.; Kanagawa, K.; Negoro, S.; Okada, H.
J. Bacteriol. 171, 3187-3191, 1989

A;Title: High homology between 6-aminohexanoate-cyclic-dimer hydrolases of Flavobacterium
A;Reference number: A44761; MUID:89255082; PMID:2722746
A;Accession: B44761
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <TSU>
A;Cross-references: GB:M26953; NID:g148710; PIDN:AAA24929.1; PID:g148711
C;Genetics:
A;Start codon: GTG
C;Superfamily: indoleacetamide hydrolase
C;Keywords: hydrolase

Query Match 82.9%; Score 29; DB 2; Length 493;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
Db 421 MLFTVP 427

RESULT 21
A44761
6-aminohexanoate-cyclic-dimer hydrolase (EC 3.5.2.12) - Pseudomonas sp. (strain NK87)
C;Species: Pseudomonas sp.
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Oct-1999
R;Accession: A44761
R;Tsuchiya, K.; Fukuyama, S.; Kanzaki, N.; Kanagawa, K.; Negoro, S.; Okada, H.
J. Bacteriol. 171, 3187-3191, 1989
A;Title: High homology between 6-aminohexanoate-cyclic-dimer hydrolases of Flavobacterium
A;Reference number: A44761; MUID:89255082; PMID:2722746
A;Accession: A44761
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <TSU>
A;Cross-references: GB:M26952; NID:g151402; PIDN:AAA25908.1; PID:g151403
C;Genetics:
A;Start codon: GTG
C;Superfamily: indoleacetamide hydrolase
C;Keywords: hydrolase

Query Match 82.9%; Score 29; DB 2; Length 493;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
Db 421 MLFTVP 427

RESULT 22
G75151
trk potassium uptake system protein (trkH) PAB0248 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G75151
R;anonymous, Genoscope
Submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A;Reference number: A75001
A;Accession: G75151
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <KAW>
A;Cross-references: UNIPROT:Q9V1Q5; GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB4929
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: trkH; PAB0248
C;Superfamily: potassium uptake protein trkG

Query Match 82.9%; Score 29; DB 2; Length 493;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 8
Db 286 LLAATIP 293

RESULT 23

G75267

ABC transporter, permease protein, CystW family - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: G75267
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; J.
M.; Shen, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radiorasiatant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75267
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-515 <WHI>
A;Cross-references: UNIPROT:Q9RRK1; GB:AE002078; GB:AE000513; NID:g6460306; PIDN:AAF1202
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2488
A;Map position: 1
C;Superfamily: sfuB protein

Query Match 82.9%; Score 29; DB 2; Length 515;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 8
Db 379 VLAAATIP 386

RESULT 24

D87753

protein C43E11.6 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D87753
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: D87753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-721 <STO>
A;Cross-references: UNIPROT:P91146; GB:chr_I; PIDN:AAB37620.1; PID:g1703566; GSPDB:GN000
C;Genetics:
A;Gene: C43E11.6
A;Map position: 1

Query Match 82.9%; Score 29; DB 2; Length 721;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 8
Db 714 VIAATIP 721

RESULT 25

D72079

conserved hypothetical protein CP0319 [imported] - Chlamydomophila pneumoniae (strains CWL
N;Alternate names: ct283 hypothetical protein
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
A:Accession: D72079; G81588
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72079
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-729 <ARN>
A:Cross-references: UNIPROT:Q9Z8A9; GB:AE001626; GB:AE001363; NID:g4376708; PIDN:AAD1857
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: G81588
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-729 <REA>
A:Cross-references: GB:AE002195; GB:AE002161; NID:g7189246; PIDN:AAF38174.1; PID:g718924
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CPn0434; CP0319
C:Superfamily: Chlamydia trachomatis hypothetical protein CT283
Query Match 82.9%; Score 29; DB 2; Length 729;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LAXTIPM 8
Db 441 LAATVPM 447
RESULT 26
H86544
CT283 hypothetical protein [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H86544
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: H86544
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-729 <STO>
A:Cross-references: UNIPROT:Q9Z8A9; GB:BA000008; NID:g8978806; PIDN:BAA98642.1; GSPDB:GN
A:Experimental source: strain J138
C:Genetics:
A:Gene: CPJ0434
C:Superfamily: Chlamydia trachomatis hypothetical protein CT283
Query Match 82.9%; Score 29; DB 2; Length 729;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LAXTIPM 8
Db 441 LAATVPM 447
RESULT 27
T24611
hypothetical protein K08H10.2b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24611
R:Wilkinson, J.
submitted to the EMBL Data Library, June 1996

A:Reference number: Z19912
A:Accession: T24611
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-24 <WIL>
A:Cross-references: UNIPROT:Q9XVV2; EMBL:Z73975; PIDN:CAA98278.1; GSPDB:GN000023; CESP:K0
A:Experimental source: clone T06E8
C:Genetics:
A:Gene: CESP:K08H10.2b
A:Map position: 5
A:Introns: 23/2
Query Match 80.0%; Score 28; DB 2; Length 24;
Best Local Similarity 62.5%; Pred. No. 10;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MLAXTIPM 8
Db 1 MLRTLEPM 8
RESULT 28
S72786
hypothetical protein B1549_F2_59 - Mycobacterium leprae
C:Species: Mycobacterium leprae
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S72786
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B1549.
A:Reference number: S72582
A:Accession: S72786
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-119 <SMI>
A:Cross-references: UNIPROT:P54133; EMBL:U00014; NID:g466903; PIDN:AAA50901.1; PID:g4669
Query Match 80.0%; Score 28; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLAXTIPM 8
Db 88 LLAGTVPL 95
RESULT 29
E86662
phosphonate ABC transporter permease protein phnB [imported] - Lactococcus lactis subsp.
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86662
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: E86662
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <STO>
A:Cross-references: UNIPROT:Q9C1Q5; GB:AE005176; PID:g12723165; PIDN:AAK04399.1; GSPDB:G
A:Experimental source: strain Ill1403
C:Genetics:
A:Gene: phnB
C:Superfamily: phnE protein
Query Match 80.0%; Score 28; DB 2; Length 270;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLAXTIP 7
Db 123 MLARTLP 129

Wed Feb 9 06:11:22 2005

```

C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S76464
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76464
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-465 <KAN>
A;Cross-references: UNIPROT:P74491; EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BAAL859
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match      80.0%; Score 28; DB 2; Length 465;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAXTIP 7
      |||:|
Db      65 MIANTIP 71

RESULT 33
S46752
hypothetical protein YHR159w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: S46752
R;Macri, C.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of S. cerevisiae cosmid 9666.
A;Reference number: S46752
A;Accession: S46752
A;Molecule type: DNA
A;Residues: 1-504 <MAC>
A;Cross-references: UNIPROT:P38854; EMBL:U10397; NID:g500647; PID:g500648; GSPDB:GN00008
C;Genetics:
A;Gene: MIFS:YHR159w
A;Cross-references: SGD:S0001202
A;Map position: 8R

Query Match      80.0%; Score 28; DB 2; Length 504;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAXTIP 7
      |||:|
Db      165 MIANTIP 171

RESULT 34
H84141
L-lactate permease lctp [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H84141
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84141
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-524 <STO>
A;Cross-references: UNIPROT:Q9K5Z9; GB:AF001520; GB:BA000004; NID:g10176401; PIDN:BA5076
A;Experimental source: strain C-125
C;Genetics:
A;Gene: lctp
C;Superfamily: L-lactate permease

```

```

RESULT 30
D95882
probable efflux protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid ps
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: D95882
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: D95882
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-407 <KUR>
A;Cross-references: UNIPROT:Q92WL5; GB:AL591985; PIDN:CAC48724.1; PID:g15140197; GSPDB:G
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
C;Genetics:
A;Contents: annotation
A;Gene: Smb20338
A;Genome: plasmid

Query Match      80.0%; Score 28; DB 2; Length 407;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAXTIP 8
      |||:|
Db      295 LLADIVPL 302

RESULT 31
A87305
Na+/H+ antiporter [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87305
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87305
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <STO>
A;Cross-references: UNIPROT:Q9AAZ2; GB:AE005673; NID:g13421621; PIDN:AAK22437.1; GSPDB:G
C;Genetics:
A;Gene: CC0450
C;Superfamily: Na+/H+-exchanging protein nhax

Query Match      80.0%; Score 28; DB 2; Length 413;
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAXTIP 8
      |||:|
Db      230 VLAMTIP 237

RESULT 32
S76464
hypothetical protein - Synechocystis sp. (strain PCC 6803)

```



```
Query Match      80.0%; Score 28; DB 2; Length 524;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAXTIPM 8
DB      195 MTALTLP 202

RESULT 35
T30629
Hypothetical protein 27L - Molluscum contagiosum virus 1
N:Alternate names: MC027L
C:Species: Molluscum contagiosum virus 1
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T30629
R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science 273, 813-816, 1996
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
A:Reference number: Z20876; MUID:96325459; PMID:8670425
A:Accession: T30629
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-534 <SEN>
A:Cross-references: UNIPROT:Q98195; EMBL:U60315; NID:g1491943; PIDN:AAC55155.1; PID:g149
C:Genetics:
A:Note: MC027L
C:Superfamily: Molluscum contagiosum virus 1 hypothetical protein 27L

Query Match      80.0%; Score 28; DB 2; Length 534;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAXTIP 7
DB      305 MLAGTLP 311

RESULT 36
A12327
Hypothetical protein alr4176 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: A12327
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: A12327
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-545 <KUR>
A:Cross-references: UNIPROT:Q8YPL8; GB:BA000019; PIDN:BA875875.1; PID:g17133311; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4176

Query Match      80.0%; Score 28; DB 2; Length 545;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAXTIP 7
DB      178 MNAFTLP 184

RESULT 37
H75577
probable drug transport protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: H75577
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75577
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-698 <WHI>
A:Cross-references: UNIPROT:Q9RYR1; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF12339
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0247
A:Map position: 2

Query Match      80.0%; Score 28; DB 2; Length 698;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 MLAXTIPM 8
DB      237 ILATTIPL 244

RESULT 38
JC4837
hypoxia-inducible factor 1 alpha - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
C:Accession: JC4837
R:Wenger, R.H.; Rolfs, A.; Marti, H.H.; Guenet, J.L.; Gassmann, M.
Biochem. Biophys. Res. Commun. 223, 54-59, 1996
A:Title: Nucleotide sequence, chromosomal assignment and mRNA expression of mouse hypoxi
A:Reference number: JC4837; MUID:96254028; PMID:8660378
A:Accession: JC4837
A:Molecule type: mRNA
A:Residues: 1-810 <WEN>
A:Cross-references: UNIPROT:Q61221; EMBL:X95580; NID:g1430864; PIDN:CAA64833.1; PID:g437
C:Comment: This factor is involved in the oxygen-regulated transcription of several gene
C:Genetics:
A:Gene: Hif1alpha
A:Map position: 12
C:Keywords: transcription factor
F:5-58/Region: helix-loop-helix #status predicted

Query Match      80.0%; Score 28; DB 2; Length 810;
Best Local Similarity 75.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MLAXTIPM 8
DB      548 MLAPYIPM 555

RESULT 39
JC7619
hypoxia-inducible factor 1 alpha - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7619
R:Takahashi, T.; Sugishita, Y.; Nojiri, T.; Shinizu, T.; Yao, A.; Kinugawa, K.; Harada,
Biochem. Biophys. Res. Commun. 281, 1057-1062, 2001
A:Title: Cloning of hypoxia-inducible factor 1 alpha cDNA from chick embryonic ventricul
A:Reference number: JC7619; MUID:21134360; PMID:11237772
A:Contents: Embryonic ventricular myocytes
A:Accession: JC7619
A:Molecule type: mRNA
A:Residues: 1-811 <TAK>
A:Cross-references: DDBJ:AB013746
C:Comment: This factor belongs to the basic helix-loop-helix-Per/ARNT/Sim (HLH-PAS) fami
```

eostasis of cells, and in redox stimuli.

C:Genetics:

A:Gene: hif-1alpha
C:Keywords: embryo; transcription factor
F:106-156/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS1>
F:249-299/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS2>
F:762-811/Domain: conserved carboxy-terminal transactivation element #status predicted <
F:767-768/Region: conserved dileucine repeat, important for oxygen-dependent degradation

Query Match 80.0%; Score 28; DB 2; Length 811;
Best Local Similarity 75.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
|||: |||
Db 559 MLAPYIPM 566

RESULT 40

JC5809
hypoxia-inducible factor 1 alpha - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 26-Aug-1999
C:Accession: JC5809
R:Ladoux, A.; Frelin, C.
Biochem. Biophys. Res. Commun. 240, 552-556, 1997
A:Title: Cardiac expressions of HIF-1 alpha and HLF/EPAS, two basic loop helix/PAS domain
A:Reference number: JC5809; MUID:98063274; PMID:9398602
A:Accession: JC5809
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-813 <LAP>
C:Comment: This protein associates to the aryl hydrocarbon receptor nuclear translocator
as erythropoietin, vascular endothelial growth factor, the GLUT1 glucose transporter, and
F:6-144/Region: basic helix-loop-helix #status predicted

Query Match 80.0%; Score 28; DB 2; Length 813;
Best Local Similarity 75.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
|||: |||
Db 548 MLAPYIPM 555

RESULT 41

I38972
hypoxia-inducible factor 1 alpha - human
N:Alternate names: ARNT interacting protein
C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: I38972; G01875
R:Wang, G.L.; Jiang, B.H.; Rue, E.A.; Semenza, G.L.
Proc. Natl. Acad. Sci. U.S.A. 92, 5510-5514, 1995
A:Title: Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulate
A:Reference number: I38972; MUID:95296340; PMID:7539918
A:Accession: I38972
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-826 <RES>
A:Cross-references: UNIPROT:Q16665; EMBL:U22431; NID:g881345; PIDN:AAC50152.1; PID:g8813
A:Note: Parts of this sequence were confirmed by peptide sequencing
R:Hogenesch, J.B.; Chan, W.K.; Carver, L.A.; Bradfield, C.A.
submitted to the EMBL Data Library, June 1995

A:Reference number: H00692
A:Accession: G01875
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-826 <HOG>
A:Cross-references: EMBL:U29165; NID:g1144012; PIDN:AAC51210.1; PID:g1144013
C:Genetics:
A:Gene: GDB:HIF1A
A:Cross-references: GDB:512229

A:Map position: 14q21-14q24
C:Keywords: heterodimer

Query Match 80.0%; Score 28; DB 2; Length 826;
Best Local Similarity 75.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
|||: |||
Db 561 MLAPYIPM 568

RESULT 42

S51293
probable membrane protein YNL321w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein N0339
C:Species: Saccharomyces cerevisiae
C:Date: 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C:Accession: S51293; S59562; S63302
R:Nicaud, J.J.

submitted to the EMBL Data Library, January 1995
A:Description: Sequence analysis of a 13.9 Kb fragment of yeast chromosome XIV identifies
A:Reference number: S51285
A:Accession: S51293
A:Molecule type: DNA
A:Residues: 1-908 <NIC>
A:Cross-references: UNIPROT:P42839; EMBL:Z46259; NID:g633655; PID:g633664
R:Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
Yeast 11, 1077-1085, 1995
A:Title: Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV identifies si;
A:Reference number: S59562; MUID:96076632; PMID:7502583
A:Accession: S59562
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-908 <MAW>
A:Cross-references: EMBL:Z46259; NID:g633655; PIDN:CAA86376.1; PID:g633664
R:Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
submitted to the Protein Sequence Database, April 1996

A:Reference number: S63287
A:Accession: S63302
A:Molecule type: DNA
A:Residues: 1-908 <MAW>

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
R:Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63287
A:Accession: S63302
A:Molecule type: DNA
A:Residues: 1-908 <MAW>
A:Cross-references: EMBL:Z71597; NID:g1302433; PID:e239767; PID:g1302434; MIPS:YNL321w
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0005265

A:Map position: 14L
C:Keywords: transmembrane protein
F:250-266/Domain: transmembrane #status predicted <TM1>
F:413-429/Domain: transmembrane #status predicted <TM2>
F:496-512/Domain: transmembrane #status predicted <TM3>
F:531-547/Domain: transmembrane #status predicted <TM4>
F:561-577/Domain: transmembrane #status predicted <TM5>
F:595-611/Domain: transmembrane #status predicted <TM6>
F:631-647/Domain: transmembrane #status predicted <TM7>
F:689-705/Domain: transmembrane #status predicted <TM8>
F:748-764/Domain: transmembrane #status predicted <TM9>
F:820-836/Domain: transmembrane #status predicted <TM10>
F:886-902/Domain: transmembrane #status predicted <TM11>

Query Match 80.0%; Score 28; DB 2; Length 908;
Best Local Similarity 62.5%; Pred. No. 4.6e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
|||: |||
Db 424 LLVFTIPM 431

RESULT 43

B98341
hypothetical protein AGR_L_3349 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: B98341
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B98341
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-57 <KUR>
A:Cross-references: UNIPROT:Q8U4U6; GB:AE007870; PIDN:AAK90252.1; PID:g15160269; GSPDB:G000000000
C:Genetics:
A:Gene: AGR_L3349
A:Map position: linear chromosome

Query Match 77.1%; Score 27; DB 2; Length 57;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|:|:|:
Db 1 MVAVTVPV 8

RESULT 44
G83514
conserved hypothetical protein PA1058 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: G83514
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Beman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G83514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-89 <STO>
A:Cross-references: UNIPROT:Q9IAR6; GB:AE004537; GB:AE004091; NID:g9946960; PIDN:AAG0444
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1058

Query Match 77.1%; Score 27; DB 2; Length 89;
Best Local Similarity 62.5%; Pred. No. 69;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|:|:|:
Db 1 MLAYVIPL 8

RESULT 45
I52333
Gl phase-specific gene - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: I52333
R:Wu, G.; Su, S.; Kung, T.Y.; Bird, R.C.
Biochem. Cell Biol. 71, 372-380, 1993
A:Title: Molecular cloning of Gl phase mRNAs from a subtractive Gl phase cDNA library.
A:Reference number: I52333; MUID:94168720; PMID:8123253
A:Accession: I52333
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-109 <RES>
A:Cross-references: UNIPROT:Q16164; GB:S70622; NID:g545857; PIDN:AAB30172.1; PID:g545858
A:Experimental source: HeLa S3 cells

Query Match 77.1%; Score 27; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 86;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
:|:|:|:
Db 82 LLACTIP 88

RESULT 46
AB1127
hypothetical protein lmo0417 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1127
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1127
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-136 <GLA>
A:Cross-references: UNIPROT:Q8Y9V3; GB:NC_003210; PIDN:CAC98496.1; PID:g16409794; GSPDB:G000000000
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0417

Query Match 77.1%; Score 27; DB 2; Length 136;
Best Local Similarity 62.5%; Pred. No. 11e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|:|:|:
Db 89 VLALTIPI 96

RESULT 47
AG1487
hypothetical protein lin0438 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG1487
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1487
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-136 <GLA>
A:Cross-references: UNIPROT:Q92EM0; GB:ALU592022; PIDN:CAC95671.1; PID:g16412867; GSPDB:G000000000
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0438

Query Match 77.1%; Score 27; DB 2; Length 136;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
:|:|:|:
Db 89 VLALTIPI 96

RESULT 48

F95233
universal stress protein family [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95233
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
son, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95233
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-150 <KUR>
A;Cross-references: UNIPROT:Q97NM4; GB:AE005672; PIDN:AAK76063.1; PID:g14973504; GSPDB:C
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1996

Query Match 77.1%; Score 27; DB 2; Length 150;
Best Local Similarity 71.4%; Pred. NO. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
:|:|:|
Db 96 LLARTIP 102

RESULT 49

H98097
conserved hypothetical protein spr1810 [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: H98097
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
y, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: H98097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-171 <KUR>
A;Cross-references: UNIPROT:Q8DNC6; GB:AE007317; PIDN:AAL00613.1; PID:g15459496; GSPDB:C
C;Genetics:
A;Gene: spr1810

Query Match 77.1%; Score 27; DB 2; Length 171;
Best Local Similarity 71.4%; Pred. NO. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
:|:|:|
Db 117 LLARTIP 123

RESULT 50

G70720
probable pgsA2 protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: G70720
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: G70720

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-209 <COL>

A;Cross-references: UNIPROT:Q50611; GB:278020; GB:AL123456; NID:g3261625; PIDN:CAB01479.

A;Experimental source: strain H37RV

C;Genetics:

A;Gene: pgsA2

Query Match 77.1%; Score 27; DB 2; Length 209;
Best Local Similarity 50.0%; Pred. NO. 1.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 8
:|:|:|
Db 108 LLAGTULP 115

Search completed: February 8, 2005, 20:32:26
Job time : 20.2807 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 19:39:04 ; Search time 39.7193 Seconds
(without alignments)
103.140 Million cell updates/sec

Title: US-10-032-361-4

Perfect score: 35

Sequence: 1 MLAXTIPM 8

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 65 summaries

Database : UniProt_03.*

1: uniprot_eprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	298	Q6N0X7	Q6N0X7 rhodopsudo
2	35	100.0	553	Q9ATN6	Q9atn6 capsicum an
3	34	97.1	145	Q82622	Q82622 streptomyc
4	32	91.4	320	Q87LN5	Q87ln5 vibrio para
5	32	91.4	321	Q9KPA8	Q9kpa8 vibrio chol
6	32	91.4	462	Q970L9	Q970l9 sulfobolus
7	31	88.6	145	Q93IV3	Q93iv3 streptomyc
8	31	88.6	462	Q8WV53	Q8wv53 homo sapien
9	31	88.6	505	Q9V9Y0	Q9v9y0 drosophila
10	31	88.6	588	Q6N3P7	Q6n3p7 rhodopsudo
11	31	88.6	678	1	CICL_RABIT
12	31	88.6	687	1	CICK_HUMAN
13	31	88.6	687	1	CICK_MOUSE
14	31	88.6	687	1	CICK_RABIT
15	31	88.6	687	1	CICK_RAT
16	31	88.6	687	1	CICK_HUMAN
17	31	88.6	687	2	Q66HN9
18	30	85.7	222	Q88VW1	Q88vw1 lactobacill
19	30	85.7	240	Q87YL4	Q87yl4 pseudomonas
20	30	85.7	300	Q93RD6	Q93rd6 brevibacter
21	30	85.7	356	Q8KE77	Q8ke77 chlorobium
22	30	85.7	459	2	Q7ON35
23	30	85.7	469	2	Q9NL14
24	30	85.7	631	2	Q72PA5
25	30	85.7	631	2	Q87F42
26	30	85.7	650	2	Q97RE5
27	30	85.7	650	2	Q8DQ95
28	30	85.7	690	2	Q6AR00
29	30	85.7	739	2	Q9PAL6
30	30	85.7	739	2	Q9PFE6
31	30	85.7	765	2	Q64A99

32	30	85.7	926	2	Q7MGB4	Q7mgb4 vibrio vuln
33	30	85.7	926	2	Q8D4R0	Q8d4r0 vibrio vuln
34	30	85.7	2471	2	Q7UJDO	Q7ujdo rhodopirell
35	29	82.9	120	1	YD42 MYCTU	Q11012 mycobacteri
36	29	82.9	145	2	Q92KE4	Q92ke4 rhizobium m
37	29	82.9	207	1	SOMA LABRO	Q9w6j7 labeo rohlt
38	29	82.9	210	2	Q90W30	Q90w30 cirrhinus m
39	29	82.9	210	2	Q90WV7	Q90wv7 catla catia
40	29	82.9	211	2	Q9W798	Q9w798 catla catia
41	29	82.9	237	2	Q8RX02	Q8rx02 arabidopsis
42	29	82.9	243	2	Q9YFO2	Q9yfq2 aeropyrum p
43	29	82.9	246	2	Q7NSL6	Q7nsl6 chromobacte
44	29	82.9	248	2	Q987H1	Q987h1 rhizobium l
45	29	82.9	248	2	Q6FC37	Q6fc37 acinetobact
46	29	82.9	255	1	RS2 AGRT5	Q8ufm3 agrobacteri
47	29	82.9	255	1	RS2 RHIME	Q92q55 rhizobium m
48	29	82.9	261	2	Q9SUNG	Q9suw6 arabidopsis
49	29	82.9	331	2	Q89KP3	Q89kp3 bradyrhizob
50	29	82.9	344	2	Q6WBX5	Q6wbx5 phrynosoma
51	29	82.9	353	2	Q7PLW3	Q7plw3 drosophila
52	29	82.9	378	2	Q64NJ9	Q64nj9 bacteroides
53	29	82.9	405	2	Q9RVU6	Q9rvu6 deinococcus
54	29	82.9	430	2	Q8L8L6	Q8l8l6 arabidopsis
55	29	82.9	430	2	Q9C8C9	Q9c8c9 arabidopsis
56	29	82.9	448	2	Q889G5	Q889g5 pseudomonas
57	29	82.9	462	2	Q9A6Q6	Q9a6q6 caulobacter
58	29	82.9	463	2	Q8Y4H3	Q8y4h3 listeria mo
59	29	82.9	463	2	Q928C3	Q928c3 listeria in
60	29	82.9	463	2	Q71WV8	Q71wv8 listeria mo
61	29	82.9	492	1	NYLA_FLASK	P13397 flavobacter
62	29	82.9	492	1	NYLA_PSES8	P13398 pseudomonas
63	29	82.9	493	2	Q9V1Q5	Q9v1q5 pyrococcus
64	29	82.9	515	2	Q9RRK1	Q9rrk1 deinococcus
65	29	82.9	561	2	Q95Q64	Q95q64 caenorhabdi

ALIGNMENTS

RESULT 1

ID	Q6N0X7	PRELIMINARY;	PRT;	298 AA.
AC	Q6N0X7;			
DT	05-JUL-2004 (TrEMBLrel. 27, Created)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	Short-chain dehydrogenase (EC 1.1.1.100).			
GN	Name=fixR2; OrderedLocusNames=RPA4633;			
OS	Rhodopseudomonas palustris.			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OC	Bradyrhizobiaceae; Rhodopseudomonas.			
OX	[1] NCBI_TaxID=1076;			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=CGA009 / ATCC BAA-98;			
RC	PubMed=14704707; DOI=10.1038/nbt923;			
RA	Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,			
RA	Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,			
RA	Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,			
RA	Harrison F.H., Gibson J., Harwood C.S.;			
RT	"Complete genome sequence of the metabolically versatile			
RL	photosynthetic bacterium Rhodopseudomonas palustris.";			
RL	Nat. Biotechnol. 22:55-61(2004).			
CC	-1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases			
CC	(GDR) family.			
CC	EMBL; BX572607; CAE30073.1; -.			
DR	HSSP; P25529; 1AHI.			
DR	GO; GO:0004316; F:3-oxoacyl-[acyl-carrier protein] reductase . . . ; IEA.			
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	InterPro; IPR002198; Adh_short.			
DR	InterPro; IPR002347; Adh_short_C2.			
DR	Pfam; PF00106; adh_short; 1.			

```

DR PRINTS; PR00081; GHRDR.
DR PRINTS; PR00080; SORFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 298 AA; 31846 MW; 0FCB8D367A162C4 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 298;
Best Local Similarity 87.5%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 254 MLAPTIPM 261
|||||

RESULT 2
ID Q9ATN6 PRELIMINARY; PRT; 553 AA.
AC Q9ATN6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sesquiterpene cyclase.
GN Name=PSC2;
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
[1]
RN SEQUENCE FROM N.A.
RA Kim J.-B., Ha S.-H., Lee S.-K., Lee M., Lee J.-Y.;
RL Submitted (DSC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326117; AAK15641.1; -.
DR HSBP; Q40577; SEAU.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR008129; Terpenoid synth. C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc_Teroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth_C; 1.
SQ SEQUENCE 553 AA; 63946 MW; E774122F96A2394F CRC64;

Query Match 100.0%; Score 35; DB 2; Length 553;
Best Local Similarity 87.5%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 294 MLAPTIPM 301
|||||

RESULT 3
Q82622
ID Q82622 PRELIMINARY; PRT; 145 AA.
AC Q82622;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=SAV7033;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572949; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005048; BAC74744.1; -.
KW Complete proteome.
SQ SEQUENCE 145 AA; 15089 MW; 1415406F5247185B CRC64;

Query Match 97.1%; Score 34; DB 2; Length 145;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 6 MLAPTIPM 13
|||||

RESULT 4
Q87LN5 PRELIMINARY; PRT; 320 AA.
ID Q87LN5
AC Q87LN5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sigma-E factor regulatory protein RseB.
GN OrderedLocNames=VP2576;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005082; BAC60839.1; -.
DR InterPro; IPR005588; MucB_RseB.
DR Pfam; PF03888; MucB_RseB_1.
KW Complete proteome.
SQ SEQUENCE 320 AA; 36042 MW; F8C07A27C63BF2F9 CRC64;

Query Match 91.4%; Score 32; DB 2; Length 320;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 106 MVAPTIPM 113
|||||

RESULT 5
Q9KPA8 PRELIMINARY; PRT; 321 AA.
ID Q9KPA8
AC Q9KPA8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sigma-E factor regulatory protein RseB.

```

OrderedLocusNames=VC2465;
 GN Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
 RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
 RA Sellers P., McDonald L.A., Uitterback T.R., Fleischmann R.D.,
 RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
 RA Mekalanos J.J., Venter J.C., Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae";
 RL Nature 406:477-483(2000).
 DR EMBL; AE004316; AAF95607.1; -. X
 DR PIR; F82073; F82073.
 DR TIGR; VC2465; -.
 DR InterPro; IPR005588; MucB_Rseb.
 DR Pfam; PF03888; MucB_Rseb; 1.
 KW Complete proteome.
 SQ SEQUENCE 321 AA; 36165 MW; A0E94994F41986E1 CRC64;
 Query Match 91.4%; Score 32; DB 2; Length 321;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLAXTIPM 8
 Db 106 MVADTIPM 113
 RESULT 6
 Q970L9 PRELIMINARY; PRT; 462 AA.
 AC Q970L9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Hypothetical protein ST1578.
 DE Hypothetical protein ST1578.
 GN OrderedLocusNames=ST1578;
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX MEDLINE=21456156; PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kuehida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermophilic
 Crenarchaeon, Sulfolobus tokodaii strain7";
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AP000986; BAB6654.1; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 462 AA; 51852 MW; 3B6F3D90630FFB6E CRC64;
 Query Match 91.4%; Score 32; DB 2; Length 462;
 Best Local Similarity 75.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLAXTIPM 8
 Db 380 LLALTIPM 387

RESULT 7
 Q93IV3 PRELIMINARY; PRT; 145 AA.
 AC Q93IV3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Hypothetical protein SC01317.
 DE Hypothetical protein SC01317.
 GN ORENames=SCBAC365.28c;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1302;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL939108; CAC42864.1; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 145 AA; 15121 MW; E3A5834748B315E2 CRC64;
 Query Match 88.6%; Score 31; DB 2; Length 145;
 Best Local Similarity 62.5%; Pred. No. 97;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLAXTIPM 8
 Db 6 LLAATVPM 13
 RESULT 8
 Q8W53 PRELIMINARY; PRT; 462 AA.
 AC Q8W53;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE Hypothetical protein.
 DE Hypothetical protein.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E., Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodger T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "The genome sequence of Drosophila melanogaster."; Science 287:2185-2195 (2000).	RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodger T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "The genome sequence of Drosophila melanogaster."; Science 287:2185-2195 (2000).
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	Tissue: Kidney;	RC	Tissue: Kidney;	RC	Tissue: Kidney;
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.	RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.	RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC020873; AAH20873.1; -	DR	EMBL; BC020873; AAH20873.1; -	DR	EMBL; BC020873; AAH20873.1; -
DR	GO; GO:0016020; C:membrane; IEA.	DR	GO; GO:0016020; C:membrane; IEA.	DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0005247; P:voltage-gated chloride channel activity; IEA.	DR	GO; GO:0005247; P:voltage-gated chloride channel activity; IEA.	DR	GO; GO:0005247; P:voltage-gated chloride channel activity; IEA.
DR	GO; GO:0006821; P:chloride transport; IEA.	DR	GO; GO:0006821; P:chloride transport; IEA.	DR	GO; GO:0006821; P:chloride transport; IEA.
DR	PFam; PF00654; Voltage CLC; 1.	DR	PFam; PF00654; Voltage CLC; 1.	DR	PFam; PF00654; Voltage CLC; 1.
DR	PRINTS; PR00762; CLCHANNEL.	DR	PRINTS; PR00762; CLCHANNEL.	DR	PRINTS; PR00762; CLCHANNEL.
DR	PRINTS; PR01119; CLCHANNELKDY.	DR	PRINTS; PR01119; CLCHANNELKDY.	DR	PRINTS; PR01119; CLCHANNELKDY.
KW	Hypothetical protein.	KW	Hypothetical protein.	KW	Hypothetical protein.
SQ	SEQUENCE 462 AA; 50638 MW; 57849DBC7B539C3P CRC64;	SQ	SEQUENCE 462 AA; 50638 MW; 57849DBC7B539C3P CRC64;	SQ	SEQUENCE 462 AA; 50638 MW; 57849DBC7B539C3P CRC64;
Query Match	88.6%; Score 31; DB 2; Length 462;	Query Match	88.6%; Score 31; DB 2; Length 505;	Query Match	88.6%; Score 31; DB 2; Length 505;
Best Local Similarity	75.0%; Pred. No. 2.9e+02;	Best Local Similarity	75.0%; Pred. No. 3.2e+02;	Best Local Similarity	75.0%; Pred. No. 3.2e+02;
Matches	6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	Matches	6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	Matches	6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY	1 MLAXTIPM 8	QY	1 MLAXTIPM 8	QY	1 MLAXTIPM 8
Db	286 ILAXTIPM 293	Db	286 ILAXTIPM 293	Db	286 ILAXTIPM 293
RESULT 9		RESULT 9		RESULT 9	
Q9V9Y0		Q9V9Y0		Q9V9Y0	
ID	Q9V9Y0 PRELIMINARY; PRT; 505 AA.	ID	Q9V9Y0 PRELIMINARY; PRT; 505 AA.	ID	Q9V9Y0 PRELIMINARY; PRT; 505 AA.
AC	Q9V9Y0	AC	Q9V9Y0	AC	Q9V9Y0
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	DT	01-MAY-2000 (TrEMBLrel. 13, Created)	DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAY-2000 (TrEMBLrel. 26, Last annotation update)	DT	01-MAY-2000 (TrEMBLrel. 26, Last annotation update)	DT	01-MAY-2000 (TrEMBLrel. 26, Last annotation update)
DE	CG1607-PA (CG1607-pb).	DE	CG1607-PA (CG1607-pb).	DE	CG1607-PA (CG1607-pb).
GN	ORFNames=CG1607;	GN	ORFNames=CG1607;	GN	ORFNames=CG1607;
OS	Drosophila melanogaster (Fruit fly).	OS	Drosophila melanogaster (Fruit fly).	OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.	OC	Ephydroidea; Drosophilidae; Drosophila.	OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_taxid=7227;	OX	NCBI_taxid=7227;	OX	NCBI_taxid=7227;
[1]		[1]		[1]	
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;	RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;	RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brannon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L., Abril J.F., Adaya V.A., An H.J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Butch J., Brockstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Fester C., Gabrielian A.C., Garg N.S., Gelbart W.M., Glasser K., Glisic A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brannon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L., Abril J.F., Adaya V.A., An H.J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Butch J., Brockstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Fester C., Gabrielian A.C., Garg N.S., Gelbart W.M., Glasser K., Glisic A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brannon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L., Abril J.F., Adaya V.A., An H.J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Butch J., Brockstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Fester C., Gabrielian A.C., Garg N.S., Gelbart W.M., Glasser K., Glisic A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,


```

Db 317 MLAWTIPV 324
|||||:
RESULT 10
Q6N3P7 PRELIMINARY; PRT; 588 AA.
AC Q6N3P7;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DE Putative maltoligosyltrehalose trehalohydrolase (SC 2.4.1.18).
GN OrderedLocName=RP3A3646;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris."
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572604; CAE29087.1; -
DR GO; GO:0003844; F:1,4-alpha-glucan branching enzyme activity; IEA.
DR GO; GO:0004556; F:alpha-amylase activity; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha amyl cat.
DR InterPro; IPR006589; Alp_aml cat sub.
DR InterPro; IPR004193; Glyco_hydro_13N.
DR Pfam; PF00128; Alpha-amylase; 1.
DR SMART; SM00642; Amy; 1.
DR Complete proteome.
SQ SEQUENCE 588 AA; 65631 MW; 051C0645B17CD0CF CRC64;

Query Match 88.6%; Score 31; DB 2; Length 588;
Best Local Similarity 62.5%; Pred. No. 3.7e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWTIPM 8
Db 417 LUAPTVM 424
|||||:

RESULT 11
CICL_RABIT STANDARD; PRT; 678 AA.
ID CICL_RABIT STANDARD; PRT; 678 AA.
AC P51804;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Chloride channel protein CIC-Kb (Chloride channel Kb) (CIC-K2).
GN Name=CICNB;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96163242; PubMed=8587242;
RA Zimniak L., Winters C.J., Reeves W.B., Andreoli T.E.;
RT "Cl- channels in basolateral renal medullary vesicles. X. Cloning of a
RT Cl- channel from rabbit outer medulla."
RL Kidney Int. 48:1828-1836(1995).
RC -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have

```

```

several functions including the regulation of cell volume;
membrane potential stabilization, signal transduction and
transepithelial transport. May be important in urinary
concentrating mechanisms.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- TISSUE SPECIFICITY: Expressed predominantly in the kidney.
-1- SIMILARITY: Belongs to the chloride channel (TC 1.A.11) family.
-1- SIMILARITY: Contains 2 CBS domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U36790; AAC48493.1; -
CC InterPro; IPR000644; CBS.
CC InterPro; IPR001807; Cl-channel_volt.
CC InterPro; IPR002250; Cl_channelKDY.
CC Pfam; PF00571; CBS; 1.
CC Pfam; PF00654; Voltage CLC; 1.
CC PRINTS; PR00762; CLCHANNEL.
CC SMART; SM00116; CBS; 1.
CC CBS domain; Chloride; Chloride channel; Ion transport; Ionic channel;
CC Repeat; Transmembrane; Voltage-gated channel.
CC DOMAIN 1 49 Cytoplasmic (Potential).
CC TRANSMEM 50 69 1 (Potential).
CC TRANSMEM 91 114 2 (Potential).
CC TRANSMEM 138 159 3 (Potential).
CC TRANSMEM 169 188 4 (Potential).
CC TRANSMEM 200 224 5 (Potential).
CC TRANSMEM 239 257 6 (Potential).
CC TRANSMEM 282 302 7 (Potential).
CC TRANSMEM 325 348 8 (Potential).
CC TRANSMEM 398 417 9 (Potential).
CC TRANSMEM 420 438 10 (Potential).
CC TRANSMEM 466 487 11 (Potential).
CC TRANSMEM 495 514 12 (Potential).
CC DOMAIN 515 645 Cytoplasmic (Potential).
CC TRANSMEM 646 664 13 (Potential).
CC DOMAIN 665 678 Extracellular (Potential).
CC DOMAIN 549 604 CBS 1.
CC DOMAIN 620 678 CBS 2.
CC CARBOHYD 364 364 N-linked (GlcNAc...) (Potential).
CC SEQUENCE 678 AA; 74468 MW; 62816AB2877125F2 CRC64;

Query Match 88.6%; Score 31; DB 1; Length 678;
Best Local Similarity 75.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWTIPM 8
Db 414 ILATTIPM 421
|||||:

RESULT 12
CICL_HUMAN STANDARD; PRT; 687 AA.
ID CICL_HUMAN STANDARD; PRT; 687 AA.
AC P51800; Q726D1; Q86VT1;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Chloride channel protein CIC-Ka (Chloride channel Ka) (CIC-K1).
GN Name=CICNKA;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;

```

```

RX MEDLINE=94316614; PubMed=8041726;
RA Kieferle S., Fong P., Bens M., Vandewalle A., Jentsch T.;
RT "Two highly homologous members of the CLC chloride channel family in
RL both rat and human kidney";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6943-6947(1994).
[2]
RN SEQUENCE FROM N.A., AND VARIANT PHE-315.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Datsenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
RN SEQUENCE OF 153-203 FROM N.A.
RX Schutte B.C., Malik M.I., Fingert J., Barna T.J., Stone E., Lamb F.S.;
RT "Refined chromosomal localization of six human CLCN chloride ion
RL channel genes";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
[4]
RN TISSUE SPECIFICITY.
RX PubMed=11734858; DOI=10.1038/35107099;
RA Estevez R., Boettger T., Stein V., Birkenhaeger R., Otto E.,
RA Hildebrandt F., Jentsch T.J.;
RT "Barttin is a Cl- channel beta-subunit crucial for renal Cl-
RT reabsorption and inner ear K+ secretion.";
RL Nature 414:558-561(2001).
CC -!- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC several functions including the regulation of cell volume;
CC membrane potential stabilization, signal transduction and
CC transepithelial transport. May be important in urinary
CC concentrating mechanisms.
CC -!- SUBUNIT: Interacts with BSNL. Forms heteromers with BSNL in the
CC thin ascending limb of Henle (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed predominantly in the kidney. All
CC nephron segments expressing BSNL also express CLCNK proteins.
CC -!- SIMILARITY: Belongs to the chloride channel (TC 1.A.11) family.
CC -!- SIMILARITY: Contains 2 CBS domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC ENBL; Z30643; CAA83120.1; --
CC ENBL; BC048282; AAH48282.1; --
CC ENBL; BC053869; AAH53869.1; --
CC ENBL; U93878; AAB65148.1; --
CC PIR; C57713; C57713.
CC Genew; HGNC:2026; CLCNKA.
CC MIM; 602024; --
CC GO; GO:0005887; C: integral to plasma membrane; TAS.
CC GO; GO:0005247; F: voltage-gated chloride channel activity; TAS.

```

DR GO:0007588; P:excretion; TAS.
 DR GO:0006810; P:transport; TAS.
 DR InterPro; IPR000844; CBS.
 DR InterPro; IPR001807; Cl-channel volt.
 DR InterPro; IPR002250; Cl_channelKDY.
 DR Pfam; PF00571; CBS; 2.
 DR Pfam; PF00654; Voltage CLC; 1.
 DR SMART; PR00762; CLCHANNEL.
 DR SMART; SM00116; CBS; 1.
 DR CBS domain; Chloride; Chloride channel; Ion transport; Ionic channel;
 KW Polymorphism; Repeat; Transmembrane; Voltage-gated channel.
 FT DOMAIN 1 49
 FT TRANSMEM 50 69 Cytoplasmic (Potential).
 FT TRANSMEM 91 114 2 (Potential).
 FT TRANSMEM 138 159 3 (Potential).
 FT TRANSMEM 169 188 4 (Potential).
 FT TRANSMEM 200 224 5 (Potential).
 FT TRANSMEM 239 257 6 (Potential).
 FT TRANSMEM 282 302 7 (Potential).
 FT TRANSMEM 325 348 8 (Potential).
 FT TRANSMEM 398 417 9 (Potential).
 FT TRANSMEM 420 438 10 (Potential).
 FT TRANSMEM 466 487 11 (Potential).
 FT TRANSMEM 495 514 12 (Potential).
 FT DOMAIN 515 645 Cytoplasmic (Potential).
 FT TRANSMEM 646 664 13 (Potential).
 FT DOMAIN 665 687 Extracellular (Potential).
 FT DOMAIN 543 604 CBS 1.
 FT DOMAIN 620 678 CBS 2.
 FT CARBOHYD 679 699 N-linked (GlcNAc...) (Potential).
 FT VARIANT 83 83 R -> G (in dbSNP:10927887).
 FT VARIANT 315 315 /FTID=VAR_019787.
 FT VARIANT 447 447 Y -> F (in dbSNP:12126269).
 FT CONFLICT 67 67 /FTID=VAR_019788.
 FT CONFLICT 615 615 A -> T (in dbSNP:1805152).
 FT CONFLICT 615 615 M -> I (in Ref. 2; AAH48282).
 FT SEQUENCE 687 AA; 75284 MW; E97C6928470AA460 CRC64; Missing (in Ref. 2; AAH53869).
 Query Match 88.6%; Score 31; DB 1; Length 687;
 Best Local Similarity 75.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MLAXTIPM 8
 Db 414 ILATTIPM 421
 RESULT 13
 CLCK_MOUSE STANDARD; PRT; 687 AA.
 AC Q9WUB7; Q9WUB7; Q8JZU7;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Chloride channel protein CLC-Ka (Chloride channel Ka) (CLC-K1).
 GN Name=CLCNka; Synonyms=CLCNk1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., POSSIBLE FUNCTION, AND TISSUE SPECIFICITY.
 RC TISSUE=Kidney cortex;
 RX MEDLINE=20469551; PubMed=11014860;
 RA Winters C.J., Zimnick L., Mikhailova M.V., Reeves W.B., Andreoli T.E.;
 RT "Cl- channels in basolateral TAL membranes XV. Molecular heterogeneity
 RL between cortical and medullary channels.";
 RL J. Membr. Biol. 177:221-230(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RN STRAIN=C57BL/6; TISSUE=Cochlea;
 RC

RA Nie L., Vazquez A.E., Feng W., Stice J.P., Yamoah E.N.;
 RT "Functional phenotype of inner ear-specific chloride channel ClC-K and
 its accessory subunit.";
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ Databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN:PV/N; TISSUE:Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.L., Skalska U., Smalius D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]

RP INTERACTION WITH BSND, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.

RX PubMed=11734858; DOI=10.1038/35107099;

RA Estevez R., Boettger T., Stein V., Birkenhaefer R., Otto E.,

RA Hildebrandt F., Jentsch T.J.;

RT "Barttin is a Cl- channel beta-subunit crucial for renal Cl-

RT reabsorption and inner ear K+ secretion.";

RL Nature 414:558-561(2001).

CC -!- FUNCTION: Voltage-gated chloride channel. Chloride channels have

several functions including the regulation of cell volume;

membrane potential stabilization, signal transduction and

trans epithelial transport. May be important in urinary

concentrating mechanisms. May be the basolateral chloride channel

mediating net chloride absorption in Ctrial cells.

CC -!- SUBUNIT: Interacts with BSND. Forms heteromers with BSND in the

thin ascending limb of Henle.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- TISSUE SPECIFICITY: Specifically expressed in the kidney. All

neuron segments expressing BSND also express CLCNK proteins.

CC -!- SIMILARITY: Belongs to the chloride channel (TC 1.A.11) family.

CC -!- SIMILARITY: Contains 2 CBS domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF124848; AAD21083.1; -;

CC EMBL; AY373832; AAQ81628.1; -;

CC EMBL; BC037077; AAA37077.1; -;

CC MGD; MGI:1329026; Clcnka.

CC GO; GO:0050878; P:regulation of body fluids; IMP.

CC InterPro; IPR000644; CBS.

CC InterPro; IPR001807; Cl-channel volt.

CC InterPro; IPR002250; Cl_channelKDY.

CC Pfam; PF00571; CBS; 2.

CC Pfam; PF00654; Voltage CLC; 1.

CC PRINTS; PR00762; CLCHANNEL.

CC PRINTS; PR01119; CLCHANNELKDY.

CC SMART; SM00116; CBS; 1.

CC CBS domain; Chloride; Chloride channel; Ion transport; Ionic channel;

KW Repeat; Transmembrane; Voltage-gated channel.

FT DOMAIN 1 49 Cytoplasmic (Potential).
 FT TRANSMEM 50 69 1 (Potential).
 FT 91 114 2 (Potential).
 FT TRANSMEM 138 159 3 (Potential).
 FT TRANSMEM 169 188 4 (Potential).
 FT TRANSMEM 200 224 5 (Potential).
 FT TRANSMEM 239 257 6 (Potential).
 FT TRANSMEM 282 302 7 (Potential).
 FT TRANSMEM 325 348 8 (Potential).
 FT TRANSMEM 398 417 9 (Potential).
 FT TRANSMEM 420 438 10 (Potential).
 FT TRANSMEM 466 487 11 (Potential).
 FT TRANSMEM 495 514 12 (Potential).
 FT DOMAIN 515 645 Cytoplasmic (Potential).
 FT TRANSMEM 646 664 13 (Potential).
 FT DOMAIN 665 687 Extracellular (Potential).
 FT DOMAIN 549 604 CBS 1.
 FT DOMAIN 626 628 CBS 2.
 FT CONFLICT 372 372 R -> W (in Ref. 3).
 FT CONFLICT 380 380 A -> S (in Ref. 2).
 SQ SEQUENCE 687 AA; 75668 MW; 602AD591DA62236F CRC64;

Query Match 88.6%; Score 31; DB 1; Length 687;

Best Local Similarity 75.0%; Pred. No. 4.3e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8

Db 414 ILATIPM 421

RESULT 14

CICK RABIT

ID_CICK RABIT STANDARD; PRT; 687 AA.

AC P51803;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Chloride channel protein ClC-Ka (Chloride channel Ka) (ClC-K1).

GN Name=CLCNKA;

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=96163242; PubMed=8587242;

RA Zimniak L., Winters C.J., Reeves W.B., Andreoli T.E.;

"Cl- channels in basolateral renal medullary vesicles. X. Cloning of a

RT Cl- channel from rabbit outer medulla.";

RL Kidney Int. 48:1828-1836(1995).

CC -!- FUNCTION: Voltage-gated chloride channel. Chloride channels have

several functions including the regulation of cell volume;

membrane potential stabilization, signal transduction and

trans epithelial transport. May be important in urinary

concentrating mechanisms.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- TISSUE SPECIFICITY: Expressed predominantly in the kidney.

CC -!- SIMILARITY: Belongs to the chloride channel (TC 1.A.11) family.

CC -!- SIMILARITY: Contains 2 CBS domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U36789; AAC48492.1; -;

CC InterPro; IPR000644; CBS.

CC InterPro; IPR001807; Cl-channel volt.

DR InterPro: IPR002250; Cl_channelKDY.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00654; Voltage_CLC; 1.
DR PRINTS: PR00762; CLCHANNEL.
DR SMART: SMO0116; CBS; 2.
DR KW CBS domain; Chloride; Chloride channel; Ion transport; Ionic channel;
DR KW Repeat; Transmembrane; Voltage-gated channel.
DR FT DOMAIN 1 49
DR TRANSMEM 50 69
DR TRANSMEM 91 114
DR TRANSMEM 138 159
DR TRANSMEM 169 188
DR TRANSMEM 200 224
DR TRANSMEM 239 257
DR TRANSMEM 282 302
DR TRANSMEM 325 348
DR TRANSMEM 398 417
DR TRANSMEM 420 438
DR TRANSMEM 466 487
DR TRANSMEM 495 514
DR TRANSMEM 515 645
DR TRANSMEM 646 664
DR TRANSMEM 665 687
DR TRANSMEM 549 604
DR TRANSMEM 620 678
DR TRANSMEM 364 364
DR CARBOHYD 687 AA; 75201 MW; B4BDB7A43078E28E CRC64;
DR SQ SEQUENCE 687 AA; 75201 MW; B4BDB7A43078E28E CRC64;

QY Query Match 88.6%; Score 31; DB 1; Length 687;
DB Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLXITPM 8
DB 414 ILATITPM 421

RESULT 15
CLICK_RAT STANDARD; PRT; 687 AA.
AC Q06393; P97709; PRT; 687 AA.
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Chloride channel protein ClC-Ka (Chloride channel Ka) (ClC-K1).
GN Name=Clcnka; Synonyms=Clcnk1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=93179373; PubMed=7680033;
RA Uchida S.; Sasaki S.; Furukawa T.; Hiraoka M.; Imai T.; Hirata Y.,
RA Marumo F.;
RT "Molecular cloning of a chloride channel that is regulated by
RT dehydration and expressed predominantly in kidney medulla.";
RL J. Biol. Chem. 268:3821-3824(1993).
RN [2]
RP REVISIONS.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=94308189; PubMed=8034678;
RA Uchida S.; Sasaki S.; Furukawa T.; Hiraoka M.; Imai T.; Hirata Y.,
RA Marumo F.;
RT "Molecular cloning of a chloride channel that is regulated by
RT dehydration and expressed predominantly in kidney medulla.";
RL J. Biol. Chem. 269:19192-19192(1994).
RN [3]
RP REVISIONS.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RA Uchida S.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.

RN SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RX MEDLINE=94316614; PubMed=8041726;
RA Kieferle S.; Fong P.; Bens M.; Vandewalle A.; Jentsch T.;
RT "Two highly homologous members of the ClC chloride channel family in
RT both rat and human kidney.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6943-6947(1994).
RN [5]
RP INTERACTION WITH BSND
RX PubMed=12111250; DOI=10.1007/s00424-002-0819-8;
RA Waldegger S.; Jock N.; Barth P.; Peters M.; Vitzthum H.; Wolf K.,
RA Kurtz A.; Konrad M.; Seyberth H.W.;
RT "Barttin increases surface expression and changes current properties
RT of ClC-K channels.";
RL Pflugers Arch. 444:411-418(2002).
RN [6]
RP DOWN-REGULATION BY FUSEMIDE
RX PubMed=12759757; DOI=10.1007/s00424-003-1098-8;
RA Wolf K.; Meier-Meitinger M.; Bergler T.; Castrop H.; Vitzthum H.,
RA Riegger G.A.J.; Kurtz A.; Kraemer B.K.;
RT "Parallel down-regulation of chloride channel ClC-K1 and barttin mRNA
RT in the thin ascending limb of the rat nephron by furosemide.";
RL Pflugers Arch. 446:665-671(2003).
CC -|- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC several functions including the regulation of cell volume;
CC membrane potential stabilization, signal transduction and
CC transepithelial transport. May be important in urinary
CC concentrating mechanisms.
CC -|- SUBUNIT: Interacts with BSND.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- TISSUE SPECIFICITY: Expressed predominantly in the kidney.
CC Expressed strongly in the cortical thick ascending limb and the
CC distal convoluted tubule, with minor expression in the S3 segment
CC of the proximal tubule and the cortical collecting tubule.
CC -|- INDUCTION: Regulated in parallel with BSND under furosemide
CC treatment. Decreased to half in the inner medulla under furosemide
CC treatment. In the renal cortex and outer medulla levels were weak
CC and did not change. Regulation with BSND in inner medulla is
CC limited to the thin limb; levels in collecting ducts were not
CC affected by furosemide treatment. During furosemide treatment
CC selective down-regulation with BSND in thin limb plays a role in
CC maintaining salt and water homeostasis.
CC -|- SIMILARITY: Belongs to the chloride channel (TC 1.A.11) family.
CC -|- SIMILARITY: Contains 2 CBS domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D13927; BAA03026.1; --
EMBL; Z34291; CAA84064.1; --
PIR; A57713; A57713.
RGD; 68435; Clcnk1.
InterPro: IPR000644; CBS.
InterPro: IPR001807; Cl-channel volt.
InterPro: IPR002250; Cl_channelKDY.
Pfam: PF00571; CBS; 2.
Pfam: PF00654; Voltage_CLC; 1.
PRINTS; PR00762; CLCHANNEL.
SMART; SMO0116; CBS; 2.
CBS domain; Chloride; Chloride channel; Ion transport; Ionic channel;
Repeat; Transmembrane; Voltage-gated channel.
DOMAIN 1 49
TRANSMEM 50 69
TRANSMEM 91 114
TRANSMEM 138 159
TRANSMEM 169 188
TRANSMEM 200 224

FT TRANSMEM 239 257 6 (Potential).
 FT TRANSMEM 282 302 7 (Potential).
 FT TRANSMEM 325 348 8 (Potential).
 FT TRANSMEM 398 417 9 (Potential).
 FT TRANSMEM 420 438 10 (Potential).
 FT TRANSMEM 466 487 11 (Potential).
 FT TRANSMEM 495 514 12 (Potential).
 FT DOMAIN 515 645 Cytoplasmic (Potential).
 FT TRANSMEM 646 664 13 (Potential).
 FT DOMAIN 665 687 Extracellular (Potential).
 FT DOMAIN 549 604 CBS 1.
 FT DOMAIN 620 678 CBS 2.
 FT CARBOHYD 364 364 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 286 286 I -> V (in Ref. 4).
 FT CONFLICT 534 534 W -> R (in Ref. 4).
 FT CONFLICT 608 609 AS -> TP (in Ref. 4).
 SQ SEQUENCE 687 AA; 75569 MW; 41434P07E3E6B8AD CRC64;

Query Match 88.6%; Score 31; DB 1; Length 687;
 Best Local Similarity 75.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
 Db 414 ILATIPM 421

RESULT 16
 C1CL_HUMAN STANDARD; PRT; 687 AA.
 AC P51801;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Chloride channel protein ClC-Kb (Chloride channel Kb) (ClC-K2).
 GN Name=CLCNKB;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=94316614; PubMed=8041726;
 RA Kieferle S., Fong P., Bens M., Vandewalle A., Jentsch T.;
 RT "Two highly homologous members of the ClC chloride channel family in
 RT both rat and human kidney";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6943-6947(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=96130539; PubMed=8544406;
 RA Takeuchi Y., Uchida S., Marumo F., Sasaki S.;
 RT "Cloning, tissue distribution, and intrarenal localization of ClC
 RT chloride channels in human kidney";
 RL Kidney Int. 48:1497-1503(1995).
 RN [3]
 RP SEQUENCE OF 151-203 FROM N.A.
 RA Schutte B.C., Malik M.I., Fingert J., Barna T.J., Stone E., Lamb F.S.;
 RT "Refined chromosomal localization of six human CLCN chloride ion
 RT channel genes";
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP FUNCTIONAL CHARACTERIZATION, AND TISSUE SPECIFICITY.
 RX PubMed=11734858; DOI=10.1038/35107099;
 RA Estevez R., Boettger T., Stein V., Birkenhaeger R., Otto E.,
 RA Hildebrandt F., Jentsch T.J.;
 RT "Barttin is a Cl- channel beta-subunit crucial for renal Cl-
 RT reabsorption and inner ear K+ secretion";
 RL Nature 414:558-561(2001).
 RN [5]
 RP VARIANTS BS TYPE 3 LEU-124; THR-204; ASP-349; HIS-432 AND CYS-438.
 RX MEDLINE=97467727; PubMed=9326936;

Simon D.B., Bindra R.S., Mansfield T.A., Nelson-Williams C.,
 Mendonca E., Stone R., Schurman S., Nayyar A., Alpay H., Bakaloglu A.,
 Rodriguez-Sotiano J., Morales J.M., Sanjad S.A., Taylor C.M., Pilz D.,
 Brem A., Trachtman H., Griswold W., Richard G.A., John E.,
 Lifton R.P.;
 RT "Mutations in the chloride channel gene, CLCNKB, cause Bartter's
 RT syndrome type III.";
 RL Nat. Genet. 17:171-178(1997).
 CC -|- FUNCTION: Voltage-gated chloride channel. Chloride channels have
 CC several functions including the regulation of cell volume;
 CC membrane potential stabilization, signal transduction and
 CC transepithelial transport. May be important in urinary
 CC concentrating mechanisms.
 CC -|- SUBUNIT: Interacts with BSND. Forms heteromers with BSND in the
 CC thick ascending limb of Henle and more distal segments (By
 CC similarity).
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -|- TISSUE SPECIFICITY: Expressed predominantly in the kidney.
 CC -|- DISEASE: Defects in CLCNKB are a cause of Bartter syndrome type 3
 CC (BS type 3) [MIM:607364]; also known as classic Bartter syndrome.
 CC It is an autosomal recessive form of often severe intravascular
 CC volume depletion due to renal salt-wasting associated with low
 CC blood pressure, hypokalemic alkalosis, hypercalciuria, and normal
 CC serum magnesium levels.
 CC -|- MISCELLANEOUS: Compared with CLCNKA/BSND, CLCNKB/BSND is more
 CC sensitive to pH and less responsive to Ca(2+).
 CC -|- SIMILARITY: Belongs to the chloride channel (TC 1.A.11) family.
 CC -|- SIMILARITY: Contains 2 CBS domains.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z30644; CAA83121.1; -;
 DR EMBL; S80315; AAB35898.1; -;
 DR EMBL; U93879; AAB65149.1; -;
 DR PIR; D57713; D57713.
 DR Genew; HGNC:2027; CLCNKB.
 DR MIM; 602023; -;
 DR MIM; 607364; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005247; F:voltage-gated chloride channel activity; TAS.
 DR GO; GO:0007588; P:excretion; TAS.
 DR GO; GO:0006810; P:transport; TAS.
 DR InterPro; IPR000644; CBS.
 DR InterPro; IPR001807; Cl_channel_KDy.
 DR Pfam; PF00571; CBS; 2.
 DR Pfam; PF00654; Voltage_CLC; 1.
 DR PRINTS; PR00762; CLCHANNEL.
 DR SMART; SM00116; CBS; 2.
 DR KW Bartter syndrome; CBS domain; Chloride; Chloride channel;
 KW Disease mutation; Ion transport; Ionic channel; Repeat; Transmembrane;
 KW Voltage-gated channel.
 FT DOMAIN 1 49 Cytoplasmic (Potential).
 FT TRANSMEM 50 69 1 (Potential).
 FT TRANSMEM 91 114 2 (Potential).
 FT TRANSMEM 138 159 3 (Potential).
 FT TRANSMEM 169 188 4 (Potential).
 FT TRANSMEM 200 224 5 (Potential).
 FT TRANSMEM 239 257 6 (Potential).
 FT TRANSMEM 282 302 7 (Potential).
 FT TRANSMEM 325 348 8 (Potential).
 FT TRANSMEM 398 417 9 (Potential).
 FT TRANSMEM 420 438 10 (Potential).
 FT TRANSMEM 466 487 11 (Potential).
 FT TRANSMEM 495 514 12 (Potential).
 FT DOMAIN 515 645 Cytoplasmic (Potential).
 FT TRANSMEM 646 664 13 (Potential).

FT DOMAIN 665 687 Extracellular (Potential).
 FT DOMAIN 549 604 CBS 1.
 FT DOMAIN 620 678 CBS 2.
 FT CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).
 FT VARIANT 88 88 S -> R (in dbSNP:5256).
 FT VARIANT 124 124 /FTId=VAR_014466.
 FT VARIANT 143 143 P -> L (in BS type 3).
 FT VARIANT 204 204 /FTId=VAR_001624.
 FT VARIANT 334 334 N -> H (in dbSNP:5259).
 FT VARIANT 349 349 /FTId=VAR_014467.
 FT VARIANT 432 432 A -> T (in BS type 3).
 FT VARIANT 438 438 /FTId=VAR_001625.
 FT VARIANT 562 562 V -> L (in dbSNP:5251).
 FT VARIANT 687 687 /FTId=VAR_014468.
 FT SEQUENCE 687 AA; 75387 MW; 35DDF654B3BD70A4 CRC64;
 Query Match 88.6%; Score 31; DB 1; Length 687;
 Best Local Similarity 75.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MLAXTIPM 8
 Db 414 ILAXTIPM 421
 RESULT 17
 Q66HN9 PRELIMINARY; PRT; 687 AA.
 AC Q66HN9;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Clcnk1 protein.
 GN Name=Clcnk1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshynki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.I., Garcia A.M., Gay L.J., Rulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzyzinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;

RA Director MGC Project;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC081761; AAH81761.1; -.
 DR InterPro: IPR006444; CBS.
 DR InterPro: IPR001807; Cl-channel volt.
 DR InterPro: IPR022250; Cl-channelKDY.
 DR InterPro: IPR000215; Prot_inh_serpin.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF00654; Voltage CLC; 1.
 DR PRINTS: PR00762; CLCHANNEL.
 DR PRINTS: PR01119; CLCHANNELKDY.
 DR SMART: SM00116; CBS; 2.
 DR PROSITE: PS00284; SERPIN; UNKNOWN 1.
 SQ SEQUENCE 687 AA; 75515 MW; 5D3A36752A9D58A9 CRC64;
 Query Match 88.6%; Score 31; DB 2; Length 687;
 Best Local Similarity 75.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MLAXTIPM 8
 Db 414 ILAXTIPM 421
 RESULT 18
 Q88VV1 PRELIMINARY; PRT; 222 AA.
 AC Q88VV1;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE Phosphoglycolate phosphatase (EC 3.1.3.18).
 GN Name=gph2; OrderedLocusNames=lp_1932;
 OS Lactobacillus plantarum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 8826 / WCFS1;
 RX MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;
 RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
 RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
 RA De Vos W.M., Siezen R.J.;
 RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
 DR EMBL: AL935257; CAD64320.1; -.
 DR GO: GO:0003824; F:catalytic activity; IEA.
 DR GO: GO:0008967; F:phosphoglycolate phosphatase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR005834; Dehal_like_hydro.
 DR Pfam: PF00702; Hydrolase; 1.
 KW Complete proteome.
 SQ SEQUENCE 222 AA; 24559 MW; F5343EC7A81E11F6 CRC64;
 Query Match 85.7%; Score 30; DB 2; Length 222;
 Best Local Similarity 85.7%; Pred. No. 2.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 LAXTIPM 8
 Db 13 LAXTIPM 19
 RESULT 19
 Q87YL4 PRELIMINARY; PRT; 240 AA.
 ID Q87YL4;
 AC Q87YL4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

DE Hypothetical protein.
GN OrderedLocusNames=PSPT03781;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.W., Haft D.H.,
RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AB016869; AA057250.1; -.
DR TIGR; PSPT03781; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 240 AA; 24859 MW; 736E5DF201EFC901 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 240;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLASTIPM 8
DB 1 MLASTLPL 8

RESULT 20
Q93RD6 PRELIMINARY; PRT; 300 AA.
AC Q93RD6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC membrane transporter homologue.
OS Brevibacterium fuscum var. dextranlyticum.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococineae; Brevibacteriaceae; Brevibacterium.
OX NCBI_TaxID=90748;
RN [1]
RP SEQUENCE FROM N.A.
RA Mizuno T., Mori H., Nishimoto M., Ito H., Matsui H., Kimura A.,
RA Honma M., Chiba S.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AB052295; BAB6091.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp_1;
DR PROSITE; PS50928; ABC_TMI; 1.
KW Transmembrane; Transport.
SQ SEQUENCE 300 AA; 32641 MW; 2CD85A9C9D5374F9 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 300;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLASTIP 7
DB 136 MLASTIP 142

RESULT 21
Q8KE77 PRELIMINARY; PRT; 356 AA.
AC Q8KE77;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein CT0813.
GN OrderedLocusNames=CT0813;
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobaculum.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
RA Kadane D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,
RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium."
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AB012849; AAM72049.1; -.
DR TIGR; CT0813; -.
DR InterPro; IPR005524; DUF318.
DR Pfam; PF03773; DUF318; 1.
KW Complete proteome.
SQ SEQUENCE 356 AA; 36487 MW; 570FCBC3EFD9A9C88 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 356;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLASTIPM 8
DB 228 MLASTVPM 235

RESULT 22
Q7QN35 PRELIMINARY; PRT; 459 AA.
AC Q7QN35;
DT 01-WAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AgCP11736 (Fragment).
GN Name=agCG43389; ORFNames=ENSANGG00000012109;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA801002746; EAA02096.1; -.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004794; F:threonine ammonia-lyase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.

```

DR GO; GO:0009097; P:isoleucine biosynthesis; IEA.
 DR InterPro; IPR001926; B6_enzyme_beta.
 DR InterPro; IPR000634; S/T_dehydrtse_BS.
 DR InterPro; IPR001721; ThrDh C.
 DR InterPro; IPR005787; Thr_dehydratEI.
 DR Pfam; PF00291; PALP; 1.
 DR Pfam; PF00585; Thr_dehydrat C; 2.
 DR TIGRFAMs; TIGR01124; ilva_2cterm; 1.
 DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
 KW Pyridoxal phosphate.
 FT NON_TER 459
 SQ SEQUENCE 459 AA; 50073 MW; AB8CA51046B6A90C CRC64;

Query Match 85.7%; Score 30; DB 2; Length 459;
 Best Local Similarity 85.7%; Pred. No. 4.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIP 7
 Db 331 MLAVTIP 337

RESULT 23
 Q9NL14 PRELIMINARY; PRT; 469 AA.
 ID Q9NL14
 AC Q9NL14
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE AMPTRP4b protein (Fragment).
 DE Name=anPTR4b;
 GN Branchiostoma belcheri (Amphioxus).
 OS Branchiostoma belcheri (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OX NCBI_TaxID=7741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2021925; PubMed=10754074;
 RA Ono-Koyanagi K., Suga H., Kato K., Miyata T.;
 RT "protein tyrosine phosphatases from amphioxus, hagfish, and ray;
 RT divergence of tissue-specific isoform genes in the early evolution of
 RT vertebrates."; J. Mol. Evol. 50:302-311(2000).
 RL J. Mol. Evol. 50:302-311(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ono K.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB033563; BAA95170.1; -.
 DR HSP; P18052; IPI5.
 DR GO; GO:0016787; P:hydrolase activity; IEA.
 DR GO; GO:0004725; P:protein tyrosine phosphatase activity; IEA.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR000387; TYR phosphatase.
 DR InterPro; IPR00242; TYR_PP.
 DR Pfam; PF00102; Y_phosphatase; 2.
 DR PRINTS; PR00700; ERTYPHPTASE.
 DR SMART; SM00194; PTPC; 2.
 DR PROSITE; PS00226; IF; UNKNOWN 1.
 DR PROSITE; PS00383; TYR PHOSPHATASE 1; 2.
 DR PROSITE; PS00566; TYR PHOSPHATASE 2; 2.
 DR PROSITE; PS00555; TYR_PHOSPHATASE_PTP; 2.
 KW Hydrolase.
 FT NON_TER 1
 SQ SEQUENCE 469 AA; 53631 MW; C2CBED629815317 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 469;
 Best Local Similarity 85.7%; Pred. No. 5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
 Db 40 LAETIPM 46

RESULT 24

Q72PA5 PRELIMINARY; PRT; 631 AA.
 ID Q72PA5
 AC Q72PA5
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 DE OrderedLocusNames=L1C12569;
 GN Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar
 OS Copenhagen).
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=44275;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fiocruz L1-130;
 RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
 RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
 RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
 RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
 RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorri H.,
 RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A.,
 RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
 RA Jeronimo S.M.B., Junqueira-de-Azevedo J.L.M., Kimura E.T.,
 RA Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
 RA de Oliveira R.C., Pereira G.G., Reis M.S., Schriefer A.,
 RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
 RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
 RT "Comparative genomics of two Leptospira interrogans serovars reveals
 RT novel insights into physiology and pathogenesis."; J. Bacteriol. 186:2164-2172(2004).
 RL J. Bacteriol. 186:2164-2172(2004).
 DR EMBL; AB017597; AAS71131.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR InterPro; IPR005495; YJGP_YJGQ.
 DR Pfam; PF03739; YJGP_YJGQ; 1.
 KW Complete proteome.
 SQ SEQUENCE 631 AA; 71148 MW; E0559BABB0DDE707 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 631;
 Best Local Similarity 62.5%; Pred. No. 6.7e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
 Db 184 LLANTLEPM 191

RESULT 25
 Q8F742 PRELIMINARY; PRT; 631 AA.
 ID Q8F742
 AC Q8F742
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 DE OrderedLocusNames=LAL106;
 GN Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
 RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
 RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
 RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
 RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
 RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
 RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
 RA Xu J.-G., Zhao G.-P.;
 RT "Unique physiological and pathogenic features of Leptospira

RT interorgans revealed by whole-genome sequencing. ";

RL Nature 422:888-893(2003).
 DR EMBL; AE011293; AAN48305.1; -.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR InterPro; IPR005495; YJGP_YJ9Q.
 DR Pfam; PF03739; YJGP_YJ9Q; 1.
 KW Complete proteome.
 SQ SEQUENCE 631 AA; 71122 MW; CE110BAEB8EFD0F CRC64;

Query Match 85.7%; Score 30; DB 2; Length 631;
 Best Local Similarity 62.5%; Pred. No. 6.7e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
 DB 184 LLANTLPM 191
 :||:||||

RESULT 26
 Q97RES PRELIMINARY; PRT; 650 AA.
 AC Q97RES;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE PTS system, fructose specific IIAC components.
 GN OrderedLocNames=SP0877;
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=13113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC BAA-334 / TIGR4;
 RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
 RA Tetzelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzaple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
 RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.";
 RL Science 293:498-506(2001).
 DR EMBL; AE007393; AAK75004.1; -.
 DR PIR; C95101; C95101.
 DR HSSP; P00550; 1A3A.
 DR TIGR; SP0877; -.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0008982; F: protein-N(Pi)-phosphonitidine-sugar phospho. . .; IEA.
 DR GO; GO:0005351; F: sugar porter activity; IEA.
 DR GO; GO:0009401; P: phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR002173; Pfam.
 DR InterPro; IPR003352; Pfam.
 DR InterPro; IPR003353; Pfam.
 DR InterPro; IPR004715; Pfam.
 DR InterPro; IPR002178; Pfam.
 DR InterPro; IPR006327; Pfam.
 DR Pfam; PF00359; Pfam.
 DR Pfam; PF02378; Pfam.
 DR Pfam; PF02379; Pfam.
 DR Pfam; PF001689; Pfam.
 DR TIGRFAMs; TIGR00829; TIGR.
 DR TIGRFAMs; TIGR00848; TIGR.
 DR TIGRFAMs; TIGR01427; TIGR.
 DR PROSITE; PS00583; PROSITE.
 KW Complete proteome.
 SQ SEQUENCE 650 AA; 66920 MW; 651863BE4217803E CRC64;

Query Match 85.7%; Score 30; DB 2; Length 650;
 Q6AR00

Best Local Similarity 75.0%; Pred. No. 6.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAXTIPM 8
 DB 458 MLAVNIPM 465
 :||:||||

RESULT 27
 Q8DQ95 PRELIMINARY; PRT; 650 AA.
 AC Q8DQ95;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Fructose specific-phosphotransferase system IIBC component (EC 2.7.1.69).
 GN Name=fruA; OrderedLocNames=spr0780;
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RX DOI=10.1128/JB.183.19.5709-5717.2001;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA Leblanc D.J., Lee L.N., Leikowitz E.J., Lu J., Matsushima P.,
 RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
 RA Class J.I.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL; AE008453; AAK9584.1; -.
 DR PIR; D97969; D97969.
 DR HSSP; P00550; 1A3A.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0008982; F: protein-N(Pi)-phosphonitidine-sugar phospho. . .; IEA.
 DR GO; GO:0005351; F: sugar porter activity; IEA.
 DR GO; GO:0009401; P: phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR002173; Pfam.
 DR InterPro; IPR003352; Pfam.
 DR InterPro; IPR003353; Pfam.
 DR InterPro; IPR004715; Pfam.
 DR InterPro; IPR002178; Pfam.
 DR InterPro; IPR006327; Pfam.
 DR Pfam; PF00359; Pfam.
 DR Pfam; PF02378; Pfam.
 DR Pfam; PF02379; Pfam.
 DR Pfam; PF001689; Pfam.
 DR TIGRFAMs; TIGR00829; TIGR.
 DR TIGRFAMs; TIGR00848; TIGR.
 DR TIGRFAMs; TIGR01427; TIGR.
 DR PROSITE; PS00583; PROSITE.
 KW Complete proteome.
 SQ SEQUENCE 650 AA; 66934 MW; 24E41AACDC6E67C2 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 650;
 Best Local Similarity 75.0%; Pred. No. 6.8e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLAXTIPM 8
 DB 458 MLAVNIPM 465
 :||:||||

RESULT 28
 Q6AR00

```

ID Q6AR00 PRELIMINARY; PRT; 690 AA.
AC Q6AR00;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Related to methyl-accepting chemotaxis protein.
GN OrderedLocusNames=DP0495;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobubaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Partmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL; CR522870; CAG3224.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004010; Cache.
DR InterPro; IPR004089; Chmtaxis.transd.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR010989; t-snare.
DR InterPro; IPR00727; T_SNARE.
DR Pfam; PF02743; Cache; I.
DR Pfam; PF06672; HAMP; 1.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; NA; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50192; T_SNARE; 1.
KW Complete proteome.
SQ SEQUENCE 690 AA; 73762 MW; 56076A8CCBEC0A41 CRC64;

```

```

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuranae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE004056; AAF85280.1; -.
DR PIR; B82552; B82552.
DR InterPro; IPR010090; Tape_meas_TP901.
DR InterPro; IPR007713; TMP.
DR Pfam; PF05017; TMP; 3.
DR TIGRFAMs; TIGR01760; tape_meas_TP901; 1.
KW Complete proteome.
SQ SEQUENCE 739 AA; 79374 MW; F463A70EDD19F2CB CRC64;

Query Match 85.7%; Score 30; DB 2; Length 739;
Best Local Similarity 85.7%; Pred. No. 7.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
Db 167 LARTIPM 173

RESULT 30
Q9PFE6 PRELIMINARY; PRT; 739 AA.
AC Q9PFE6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phage-related tail protein.
GN OrderedLocusNames=Xf0730;
OS Xylella fastidiosa
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuranae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA

```

```

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V. de Rosa A.J.M.,
RA de Rosa V.E. Jr. de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terezi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AF003915; AAF8350.1; -.
DR FIRM; F82769; F82769.
DR InterPro; IPR010090; Tape_meas_TP901.
DR InterPro; IPR00713; TMP.
DR Pfam; PF05017; TMP; 3.
DR TIGRPFAMs; TIGR01760; tape_meas_TP901; 1.
KW Complete proteome.
SQ SEQUENCE 739 AA; 79411 MW; 3BE29C1549BC370F CRC64;

Query Match 85.7%; Score 30; DB 2; Length 739;
Best Local Similarity 85.7%; Pred. No. 7.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 1;

QY 2 LAXTIPM 8
DB 167 LARTIPM 173

RESULT 31
Q64A99 PRELIMINARY; PRT; 765 AA.
AC Q64A99;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=G232E7.41;
OS uncultured archaeon GZfos32E7.
OC Archaea; environmental samples.
OX NCBI_TaxID=285378;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15353801;
RA Hallam S.J., Putnam N., Preston C.M., Dettler J.C., Rokhsar D.,
RA Richardson P.M., DeLong E.F.;
RT "Reverse methanogenesis: testing the hypothesis with environmental
RT genomics.";
RL Science 305:1457-1462(2004).
RN [2]
RP SEQUENCE FROM N.A.
RX Putnam N., Dettler J.C., Richardson P.M., Rokhsar D.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY14855; AAU83678.1; -.
KW Hypothetical protein.
SQ SEQUENCE 765 AA; 81842 MW; D56D24D401F52CFB CRC64;

Query Match 85.7%; Score 30; DB 2; Length 765;
Best Local Similarity 62.5%; Pred. No. 8e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 28 VLAXTIPM 35

RESULT 32
Q7MGB4 PRELIMINARY; PRT; 926 AA.
AC Q7MGB4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)

```

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycogen debranching enzyme.
GN OrderedLocusNames=VVA0055;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; AP005344; BAC96081.1; -.
DR InterPro; IPR000292; For/Nit_transpt.
DR InterPro; IPR010401; GDE_C.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR Pfam; PF06202; GDE_C; 1.
DR PROSITE; PS01005; FORMATE_NITRITE_TP_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 926 AA; 104965 MW; 8A8B86DD32F1313A CRC64;

Query Match 85.7%; Score 30; DB 2; Length 926;
Best Local Similarity 62.5%; Pred. No. 9.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 628 LMALTIM 635

RESULT 33
Q8D4R0 PRELIMINARY; PRT; 926 AA.
AC Q8D4R0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycogen debranching enzyme.
GN OrderedLocusNames=VV21229;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016812; AA008126.1; -.
DR InterPro; IPR000292; For/Nit_transpt.
DR InterPro; IPR010401; GDE_C.
DR InterPro; IPR008928; Glyco_trans_6hp.
DR Pfam; PF06202; GDE_C; 1.
DR PROSITE; PS01005; FORMATE_NITRITE_TP_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 926 AA; 104976 MW; E0CB36B86803910A CRC64;

Query Match 85.7%; Score 30; DB 2; Length 926;
Best Local Similarity 62.5%; Pred. No. 9.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
DB 628 LMALTIM 635

```

```

RESULT 34
Q7UJDO PRELIMINARY; PRT; 2471 AA.
AC Q7UJDO;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyketide synthase.
GN OrderedLocusNames=RB11975;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294154; CAD77328.1; -.
DR HSSP; Q8L3C8; 11YZ.
DR GO; GO:004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR011032; GroES_like.
DR InterPro; IPR000794; Ketoacyl synth.
DR InterPro; IPR008262; Lipase_AS.
DR InterPro; IPR006163; Phosphateth_bind.
DR InterPro; IPR000379; Ser_estrs.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00698; Acyl_transf_1.
DR Pfam; PF00106; adh_short_1.
DR Pfam; PF00107; adh_zinc_N; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; Ketoacyl-synt_C; 1.
DR Pfam; PF00550; PP-binding; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
KW Complete proteome; Transferase.
SQ SEQUENCE 2471 AA; 269205 MW; BE657A667E1733DB CRC64;

Query Match 85.7%; Score 30; DB 2; Length 2471;
Best Local Similarity 85.7%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
DB 2147 LAXTIPM 2153
ID YD42 MYCTU STANDARD; PRT; 120 AA.
AC Q11012;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Hypothetical protein Rv1342c/MT1383/Mb1377c.
GN OrderedLocusNames=Rv1342c, MT1383, Mb1377c; ORFNames=MTCY02B10.06c;

RESULT 35
YD42 MYCTU
ID YD42 MYCTU STANDARD; PRT; 120 AA.
AC Q11012;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Hypothetical protein Rv1342c/MT1383/Mb1377c.
GN OrderedLocusNames=Rv1342c, MT1383, Mb1377c; ORFNames=MTCY02B10.06c;

OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Winn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Strong, to M.leprae ML176.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; BX842576; CAB55378.1; -.
DR EMBL; AE000516; AAK45648.1; -.
DR EMBL; BX248338; CAD94238.1; -.
DR PIR; E70739; E70739.
DR TIGR; MT1383; -.
DR TuberculList; Rv1342c; -.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 24 44 Potential.
FT TRANSMEM 61 81 Potential.
FT TRANSMEM 86 106 Potential.
SQ SEQUENCE 120 AA; 13382 MW; 8D9C983C8E43EEAA CRC64;

Query Match 82.9%; Score 29; DB 1; Length 120;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
ID :::::

```

```
Db 94 LLAGTIPL 101
or send an email to license@isb-sib.ch) .
-----
CC EMBL; AF134200; AAD30540.1; -
CC HSSP; P01241; 1BP3.
DR InterPro; IPR009079; 4_helix_cytokine.
DR InterPro; IPR001400; Somatotropin.
DR Pfam; PF00103; Hormone_1; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 207 Somatotropin.
FT DISULFID 71 180 By similarity.
FT DISULFID 197 205 By similarity.
SQ SEQUENCE 207 AA; 23521 MW; 5235BA3CBFC07A28 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 207;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAGTIPM 8
Db 61 LLSKTIPTM 68

RESULT 38
Q90W30 PRELIMINARY; PRT; 210 AA.
AC Q90W30
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Growth hormone.
OS Cirrhinus mrigala.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cirrhinus.
OX NCBI_TaxID=59898;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RA Venugopal T., Pandian T.J., Mathavan S.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF140281; AAK74142.1; -.
DR HSSP; P01241; 1BP3.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR Pfam; PF00103; Hormone_1; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR ProDom; PD000010; Homeobox; 1.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
FT CHAIN 23 210 growth hormone.
SQ SEQUENCE 210 AA; 23791 MW; 1FCA6FDSA7A8498C CRC64;

Query Match 82.9%; Score 29; DB 2; Length 210;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAGTIPM 8
Db 61 LLSKTIPTM 68

RESULT 39
Q90WV7 PRELIMINARY; PRT; 210 AA.
ID Q90WV7
AC Q90WV7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
```

```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Growth hormone protein.
OS Catla catla (catla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Catla.
OX NCBI_TaxID=72446;
RN [1]
RP SEQUENCE FROM N.A.
RA Ravinder K., Majumdar K.C.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ravinder K., Pandian S.K., Majumdar K.C.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY053361; AAL14247.1; -.
DR EMBL; AY053360; AAL14246.1; -.
DR HSP; P01241; 1BP3.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR Pfam; PF00103; Hormone_1; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN 1; 1.
DR PROSITE; PS00338; SOMATOTROPIN 2; 1.
SQ SEQUENCE 210 AA; 23790 MW; 97D33B74DB68BF7 CRC64;

Query Match 82.9%; Score 29; DB 2; Length 210;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 61 LLSKTIPM 68

RESULT 40
Q9W798 PRELIMINARY; PRT; 211 AA.
AC Q9W798
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Growth hormone.
OS Catla catla (catla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Catla.
OX NCBI_TaxID=72446;
RN [1]
RP SEQUENCE FROM N.A.
RA Venugopal T., Pandian T.J., Mathavan S.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF140282; AAD39460.2; -.
DR HSP; P01241; 1BP3.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR009079; 4_helix_cytokine.
DR InterPro; IPR001400; Somatotropin.
DR Pfam; PF00103; Hormone_1; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN 1; 1.
DR PROSITE; PS00338; SOMATOTROPIN 2; 1.
FT CHAIN 23 211 growth hormone.
SQ SEQUENCE 211 AA; 23930 MW; D810E4A74063E5FD CRC64;

Query Match 82.9%; Score 29; DB 2; Length 211;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 61 LLSKTIPM 68

RESULT 40
Q9W798 PRELIMINARY; PRT; 211 AA.
AC Q9W798
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Growth hormone.
OS Catla catla (catla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Catla.
OX NCBI_TaxID=72446;
RN [1]
RP SEQUENCE FROM N.A.
RA Venugopal T., Pandian T.J., Mathavan S.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF140282; AAD39460.2; -.
DR HSP; P01241; 1BP3.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR009079; 4_helix_cytokine.
DR InterPro; IPR001400; Somatotropin.
DR Pfam; PF00103; Hormone_1; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN 1; 1.
DR PROSITE; PS00338; SOMATOTROPIN 2; 1.
FT CHAIN 23 211 growth hormone.
SQ SEQUENCE 211 AA; 23930 MW; D810E4A74063E5FD CRC64;

Query Match 82.9%; Score 29; DB 2; Length 211;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIPM 8
Db 61 LLSKTIPM 68

```

```

Db 61 LLSKTIPM 68

RESULT 41
Q8RX02 PRELIMINARY; PRT; 237 AA.
AC Q8RX02
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein At4g22530 (Fragment).
GN Name=At4g22530;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY090999; AAM14022.1; -.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR InterPro; IPR000051; SAM_bind.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 237 AA; 26326 MW; 24923A7BA520558D CRC64;

Query Match 82.9%; Score 29; DB 2; Length 237;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIP 7
Db 231 MLAGTVP 237

RESULT 42
Q9YFQ2 PRELIMINARY; PRT; 243 AA.
AC Q9YFQ2
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein APE0197.
GN OrderedLocNames=APE0197;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93310339; PubMed=10382966;
RX STRAIN=K1;
RC Kwarabavasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hosoyma A., Fukui S., Negai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RT DNA Res. 6:83-101(1999).
RL EMBL; AP000058; BAA79109.1; -.
DR PIR; C72776; C72776.
DR GO; GO:0003677; F:DNA binding; IEA.

```

DR InterPro: IPR001387; HTH 3.
 DR InterPro: IPR010982; Lambda_like_DNA.
 DR Pfam: PF01381; HTH 3; 1.
 DR SMART: SM00530; HTH XRE; 1.
 DR PROSITE: PS05093; HTH CROCI; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 243 AA; 26848 MW; 95D0E0A3AF430B0F CRC64;

Query Match 82.9%; Score 29; DB 2; Length 243;
 Best Local Similarity 71.4%; Pred. No. 4.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
 |||:|:
 DB 92 MLAGTVP 98

RESULT 43
 ID Q7NSL6 PRELIMINARY; PRT; 246 AA.
 AC Q7NSL6;
 DT 01-MAR-2004 (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Probable oxidoreductase.
 GN OrderedLocusNames=CV3406;
 OS Chromobacterium violaceum.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Chromobacterium.
 OX NCBI_TaxID=536;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 12472 / DSM 30191;
 RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
 RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
 RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
 RA Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.P.F.,
 RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
 RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
 RA Bordignon J., Brígido M.M., Brito C.A., Brocchi M., Burity H.A.,
 RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
 RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
 RA Crezynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
 RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
 RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
 RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
 RA Grattapaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
 RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
 RA Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
 RA di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,
 RA Nascimento P.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
 RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
 RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
 RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
 RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seunaz H.N.,
 RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
 RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
 RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
 RA Vettore A., Waesem R., Zaha A., Simpson A.J.G.;
 RT "The complete genome sequence of Chromobacterium violaceum reveals
 RT remarkable and exploitable bacterial adaptability."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family; AAQ61070.1; -.
 CC EMBL: AS016922; AAQ61070.1; -.
 DR HSSP: P19992; 1HDC.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002198; ADH short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam: PF00106; adh short; 1.
 DR PRINTS: PR00081; GDHRDH.
 DR PROSITE: PS00061; ADH SHORT; 1.
 KW Complete proteome; Oxidoreductase.
 SQ SEQUENCE 248 AA; 25670 MW; BF71A1913F379F28 CRC64;

Query Match 82.9%; Score 29; DB 2; Length 248;
 Best Local Similarity 71.4%; Pred. No. 4.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
 |||:|:
 DB 204 LALTVP 210

RESULT 45
 Q6FC37 PRELIMINARY; PRT; 248 AA.
 ID Q6FC37
 AC Q6FC37;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Sec-independent protein secretion pathway, component C.
 GN Name=tatC; OrderedLocusNames=ACIAD1520;
 OS Acinetobacter sp. (strain ADP1).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Acetivibrioaceae; Acetivibrio.

DR InterPro: IPR001387; HTH 3.
 DR InterPro: IPR010982; Lambda_like_DNA.
 DR Pfam: PF01381; HTH 3; 1.
 DR SMART: SM00530; HTH XRE; 1.
 DR PROSITE: PS05093; HTH CROCI; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 243 AA; 26848 MW; 95D0E0A3AF430B0F CRC64;

Query Match 82.9%; Score 29; DB 2; Length 243;
 Best Local Similarity 71.4%; Pred. No. 4.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAXTIP 7
 |||:|:
 DB 92 MLAGTVP 98

RESULT 43
 ID Q7NSL6 PRELIMINARY; PRT; 246 AA.
 AC Q7NSL6;
 DT 01-MAR-2004 (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Probable oxidoreductase.
 GN OrderedLocusNames=CV3406;
 OS Chromobacterium violaceum.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Chromobacterium.
 OX NCBI_TaxID=536;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 12472 / DSM 30191;
 RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
 RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
 RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
 RA Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.P.F.,
 RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
 RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
 RA Bordignon J., Brígido M.M., Brito C.A., Brocchi M., Burity H.A.,
 RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
 RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
 RA Crezynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
 RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
 RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
 RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
 RA Grattapaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
 RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
 RA Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
 RA di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,
 RA Nascimento P.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
 RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
 RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
 RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
 RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seunaz H.N.,
 RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
 RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
 RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
 RA Vettore A., Waesem R., Zaha A., Simpson A.J.G.;
 RT "The complete genome sequence of Chromobacterium violaceum reveals
 RT remarkable and exploitable bacterial adaptability."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family; AAQ61070.1; -.
 CC EMBL: AS016922; AAQ61070.1; -.
 DR HSSP: P19992; 1HDC.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002198; ADH short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam: PF00106; adh short; 1.
 DR PRINTS: PR00081; GDHRDH.
 DR PROSITE: PS00061; ADH SHORT; 1.
 KW Complete proteome; Oxidoreductase.
 SQ SEQUENCE 248 AA; 25670 MW; BF71A1913F379F28 CRC64;

Query Match 82.9%; Score 29; DB 2; Length 248;
 Best Local Similarity 71.4%; Pred. No. 4.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
 |||:|:
 DB 204 LALTVP 210

RESULT 45
 Q6FC37 PRELIMINARY; PRT; 248 AA.
 ID Q6FC37
 AC Q6FC37;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Sec-independent protein secretion pathway, component C.
 DR Name-tatC; OrderedLocusNames=ACIAD1520;
 GN Acinetobacter sp. (strain ADP1).
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Acetivibrioaceae; Acetivibrio.

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF009099; AAL42380.1; -.
 EMBL; AF008064; AAK87166.1; ALT_INIT.
 PIR; AF2745; AF2745.
 HAMAP; MF_00291; -.
 InterPro; IPR001865; Ribosomal_S2.
 InterPro; IPR005706; Ribosomal_S2_b/o.
 Pfam; PF00318; Ribosomal_S2; 1.
 PRINTS; PR00395; RIBOSOMALS2.
 TIGRFAMs; TIGR01011; rpsB_bact; 1.
 PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
 PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
 Complete proteome; Ribosomal protein.
 KW COMPLETE proteome; Ribosomal protein.
 SQ SEQUENCE 255 AA; 28200 MW; 0CD00217589F07C3 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 255;
 Best Local Similarity 71.4%; Pred. No. 4.7e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LAXTIPM 8
 |||:|:
 Db 45 LAQTVP 51

RESULT 47
 RS2_RHIME
 ID RS2_RHIME STANDARD; PRT; 255 AA.
 AC Q92Q55;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE 30S ribosomal protein S2.
 GN Name=rpsB; OrderedLocustNames=R01495; ORFNAMES=SMC02101;
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 "Analysis of the chromosome sequence of the legume symbiont
 Sinorhizobium meliloti strain 1021";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 RL -!- SIMILARITY: Belongs to the ribosomal protein S2P family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AL591787; CAC46074.1; -.
 HAMAP; MF_00291; -.
 InterPro; IPR001865; Ribosomal_S2.
 InterPro; IPR005706; Ribosomal_S2_b/o.
 Pfam; PF00318; Ribosomal_S2; 1.
 PRINTS; PR00395; RIBOSOMALS2.
 TIGRFAMs; TIGR01011; rpsB_bact; 1.
 PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
 PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
 Complete proteome; Ribosomal protein.
 KW COMPLETE proteome; Ribosomal protein.


```

SQ SEQUENCE 255 AA; 28048 MW; E3ECAD27D9BE906E CRC64;
Query Match 82.9%; Score 29; DB 1; Length 255;
Best Local Similarity 71.4%; Pred. No. 4.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
Db 45 LAQTVP 51

RESULT 48
Q9SUW6 PRELIMINARY; PRT; 261 AA.
AC Q9SUW6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein AT4g22530.
GN Name=AT4g22530; Synonyms=At4g22530;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RA Bevan M., Wedler H., Wambutt R., Bancroft I., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Wedler H., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Huan V.W., Lee J.M.,
RA Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
RA Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
RA Ecker J.R., Theologis A.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL033545; CAA22158.1; -
DR EMBL; AL161557; CAB9208.1; -
DR EMBL; AY150518; AAN13034.1; -
DR PIR; T05447; T05447.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR InterPro; IPR000051; SAM_bind.
KW Hypothetical protein.
SQ SEQUENCE 261 AA; 29106 MW; B04BAD92E79F745E CRC64;

Query Match 82.9%; Score 29; DB 2; Length 261;
Best Local Similarity 71.4%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLAXTIP 7
Db 255 MLAGTVP 261

RESULT 49
Q89KP3 PRELIMINARY; PRT; 331 AA.
AC Q89KP3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 30S ribosomal protein S2.
GN OrderedLocusNames=bll14861;

OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
CC -1- SIMILARITY: Belongs to the S2p family of ribosomal proteins.
DR EMBL; AP005952; BAC50126.1; -
DR GO; GO:0015935; C:small ribosomal subunit; IEA.
DR GO; GO:0003735; P:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001865; Ribosomal S2.
DR InterPro; IPR005706; Ribosomal S2 b/o.
DR Pfam; PF00318; Ribosomal S2; 1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRFAMs; TIGR01011; rpsB_bact; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Complete proteome; Ribonucleoprotein; Ribosomal protein.
SQ SEQUENCE 331 AA; 35732 MW; AD799C1BBFFA2C14 CRC64;

Query Match 82.9%; Score 29; DB 2; Length 331;
Best Local Similarity 71.4%; Pred. No. 6.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LAXTIPM 8
Db 45 LAQTVP 51

RESULT 50
Q6WBX5 PRELIMINARY; PRT; 344 AA.
AC Q6WBX5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NADH dehydrogenase subunit 2.
GN Name=ND2;
OS Phrynosoma modestum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Iguanidae; Phrynosomatinae;
OC Phrynosoma.
OX NCBI_TaxID=43612;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22801591; PubMed=12920297; DOI=10.1126/science.1084786;
RA Harmon L.J., Schulte J.A., Larson A., Losos J.B.;
RT "Tempo and mode of evolutionary radiation in iguanian lizards.";
RL Science 301:961-964(2003).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the complex I subunit 2 family.
DR EMBL; AY257484; AAP84422.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. .; IEA.
DR InterPro; IPR003917; NADH_oxred2.
DR InterPro; IPR010933; NADH_dehy_S2_C.
DR InterPro; IPR001750; Oxidored_q1.

```

DR Pfam; PF06444; NADH_dehy_S2_C; 1.
DR Pfam; PF00361; Oxidored_G1; 1.
DR PRINTS; PR01436; NADHDHGNASE2.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 344 AA; 37840 MW; E1A5ABA12A9134B1 CRC64;

Query Match 82.9%; Score 29; DB 2; Length 344;
Best Local Similarity 62.5%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLAXTIPM 8
||:|:|
Db 325 MLTLTIPM 332

Search completed: February 8, 2005, 20:09:51
Job time : 47.7193 secs

hypoxia-inducible factor 1 alpha - chicken
C:Species: Gallus gallus (chicken)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7619
R:Takahashi, T.; Sugishita, Y.; Nojiri, T.; Shimizu, T.; Yao, A.; Kinugawa, K.; Harada, Biochem. Biophys. Res. Commun. 281, 1057-1062, 2001
A>Title: Cloning of hypoxia-inducible factor 1 alpha cDNA from chick embryonic ventricle
A:Reference number: JC7619; MUID:21134360; PMID:11237772
A:Contents: Embryonic ventricular myocytes
A:Accession: JC7619
A:Molecule type: mRNA
A:Residues: 1-811 <TAK>
A:CROSS-references: DBJ:AB013746
C:Comment: This factor belongs to the basic helix-loop-helix-Per/ARNT/Sim (HLH-PAS) family of transcription factors, and in redox stimuli.
C:Genetics:
A:Gene: hif-1alpha
C:Keywords: embryo; transcription factor
F:106-156/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS1>
F:249-299/Domain: Per/ARNT/Sim, ligand binding, dimerization #status predicted <PAS2>
F:762-811/Domain: conserved carboxy-terminal transactivation element #status predicted <PAS3>
F:767-768/Region: conserved dileucine repeat, important for oxygen-dependent degradation

Query Match 62.1%; Score 95; DB 2; Length 811;
Best Local Similarity 94.7%; Pred. No. 7.4e-05;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 LDLEMLAXYIPMDDDFQL 30
|||||:|||||
Db 554 LDLEMLAPYIPMDDDFQL 572

RESULT 3
JC5809
hypoxia-inducible factor 1 alpha - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 26-Aug-1999
C:Accession: JC5809
R:Liadoux, A.; Prellin, C.
Biochem. Biophys. Res. Commun. 240, 552-556, 1997
A>Title: Cardiac expressions of HIF-1 alpha and HLF/EPAS, two basic loop helix/PAS domain proteins, in the heart of the rat.
A:Reference number: JC5809; MUID:98063274; PMID:9398602
A:Accession: JC5809
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-813 <LAD>
C:Comment: This protein associates to the aryl hydrocarbon receptor nuclear translocator as erythropoietin, vascular endothelial growth factor, the GLUT1 glucose transporter, and the p53 protein.
F:6-144/Region: basic helix-loop-helix #status predicted

Query Match 62.1%; Score 95; DB 2; Length 813;
Best Local Similarity 94.7%; Pred. No. 7.5e-05;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 LDLEMLAXYIPMDDDFQL 30
|||||:|||||
Db 543 LDLEMLAPYIPMDDDFQL 561

RESULT 4
I38972
hypoxia-inducible factor 1 alpha - human
N:Alternate names: ARNT interacting protein
C:Species: Homo sapiens (man)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: I38972; G01875
R:Wang, G.L.; Jiang, B.H.; Rue, E.A.; Semenza, G.L.
Proc. Natl. Acad. Sci. U.S.A. 92, 5510-5514, 1995
A>Title: Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by hypoxia
A:Reference number: I38972; MUID:95296340; PMID:7539918
A:Accession: I38972
A>Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-826 <RES>
A:CROSS-references: UNIPROT:Q16665; EMBL:U22431; NID:9881345; PIDN:AAC50152.1; PID:98813
A:Note: parts of this sequence were confirmed by peptide sequencing
R:Hogenesch, J.B.; Chan, W.K.; Carver, L.A.; Bradfield, C.A.
submitted to the EMBL Data Library, June 1995
A:Reference number: H00692
A:Accession: G01875
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-826 <HOG>
A:CROSS-references: EMBL:U29165; NID:g1144012; PIDN:AAC51210.1; PID:g1144013
C:Genetics:
A:Gene: GDB:HIF1A
A:CROSS-references: GDB:512229
A:Map position: 14q21-14q24
C:Keywords: heterodimer

Query Match 62.1%; Score 95; DB 2; Length 826;
Best Local Similarity 94.7%; Pred. No. 7.6e-05;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 LDLEMLAXYIPMDDDFQL 30
|||||:|||||
Db 556 LDLEMLAPYIPMDDDFQL 574

RESULT 5
JC7771
hypoxia inducible factor-3 alpha - human
C:Species: Homo sapiens (man)
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: JC7771
R:Hara, S.; Hamada, J.; Kobayashi, C.; Kondo, Y.; Imura, N.
Biochem. Biophys. Res. Commun. 287, 808-813, 2001
A>Title: Expression and characterization of hypoxia-inducible factor (HIF)-3alpha in human cells.
A:Reference number: JC7771; PMID:11573933
A:Contents: Kidney
A:Accession: JC7771
A:Molecule type: mRNA
A:Residues: 1-667 <HAR>
A:CROSS-references: UNIPROT:Q9V2N7; DBJ:AB054067
C:Comment: This protein is a heterodimeric transcription factor that belongs to the basic helix-loop-helix family of transcription factors and is involved in the regulation of hypoxia-inducible gene expression in human kidney.
C:Genetics:
A:Gene: hif-3alpha
A:Map position: 19
C:Keywords: kidney

Query Match 52.9%; Score 81; DB 2; Length 667;
Best Local Similarity 88.9%; Pred. No. 0.0047;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 LDLEMLAXYIPMDDDFQL 30
|||||:|||||
Db 483 LDLEMLAPYIPMDDDFQL 500

RESULT 6
T09384
trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate T09384)
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09384
R:Michael, N.L.; Chang, G.; d'ARCY, L.A.; Ehrenberg, P.K.; Mariani, R.; Busch, M.P.; Bilgic, A.; Viroel, J., 4228-4236, 1995
A>Title: Defective accessory genes in a human immunodeficiency virus type 1-infected cell line.
A:Reference number: Z16654; MUID:95287475; PMID:7769682
A:Accession: T09384
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-71 <MIC>
A:CROSS-references: UNIPROT:Q71926; EMBL:U24451; NID:9829440; PIDN:AAA79576.1; PID:9829440
C:Genetics:

A:Gene: tat
C:Superfamily: AIDS trans-activating transcription regulator
C:Keywords: transcription

Query Match 37.9%; Score 58; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.61; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

Qy 1 YGRKKRQRRR 11
| | | | |
Db 47 YGRKKRQRRR 57

RESULT 7

TNLJH4

trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 02-Jul-1998
C:Accession: B35523
R:Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devare,
Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986
A>Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human
A:Reference number: A94136; MUID:87041461; PMID:3490666
A:Accession: B35523
A:Molecule type: DNA
A:Residues: 1-72 <DES>
A:Cross-references: GB:M13137; NID:G326460
A:Note: the GenBank entry ADRE3AA PID:9209908 differs from the published sequence in tra
C:Genetics:
A:Gene: tat
C:Superfamily: AIDS trans-activating transcription regulator
C:Keywords: transcription regulation

Query Match 37.9%; Score 58; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.62; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

Qy 1 YGRKKRQRRR 11
| | | | |
Db 47 YGRKKRQRRR 57

RESULT 8

TNLJZR

trans-activating transcription regulator - human immunodeficiency virus Zr-6
C:Species: human immunodeficiency virus Zr-6
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C:Accession: C26192
R:Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu
Gene 52, 71-82, 1987
A>Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleoti
A:Reference number: A26192; MUID:87248097; PMID:3036660
A:Accession: C26192
A:Molecule type: DNA
A:Residues: 1-86 <SRI>
A:Cross-references: UNIPROT:P04609; GB:K03458; GB:M16322; NID:G329398; PIDN:AAA45377.1;
C:Genetics:
A:Gene: tat
A:Introns: 72/3
C:Superfamily: AIDS trans-activating transcription regulator
C:Keywords: AIDS; immunodeficiency; transcription regulation

Query Match 37.9%; Score 58; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.74; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

Qy 1 YGRKKRQRRR 11
| | | | |
Db 47 YGRKKRQRRR 57

RESULT 9

JC5591

transactivator protein - human immunodeficiency virus type 1
N:Alternate names: tat protein
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 17-Mar-1999
C:Accession: JC5591
R:Hoffmann, S.; Willbold, D.
Biochem. Biophys. Res. Commun. 235, 806-811, 1997
A>Title: A selection system to study protein-RNA interactions: Functional display of HIV
A:Reference number: JC5591; MUID:97350867; PMID:9207243
A:Accession: JC5591
A:Molecule type: protein
A:Residues: 1-86 <HO2>
C:Comment: This protein is a key regulatory protein in the viral replication cycle and b
C:Superfamily: AIDS trans-activating transcription regulator
F:22-31/Region: cysteine-rich

Query Match 37.9%; Score 58; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.74; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

Qy 1 YGRKKRQRRR 11
| | | | |
Db 47 YGRKKRQRRR 57

RESULT 10

A25700

trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 28-Sep-1987 #sequence_revision 28-Sep-1987 #text_change 09-Jul-2004
C:Accession: A25700
R:Sodroski, J.; Patarca, R.; Rosen, C.; Wong-Staal, F.; Haseltine, W.
Science 229, 74-77, 1985
A:Reference number: A25700; MUID:85244627; PMID:2990041
A:Accession: A25700
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-86 <SOD>
A:Cross-references: UNIPROT:P04610
C:Superfamily: AIDS trans-activating transcription regulator

Query Match 37.9%; Score 58; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.74; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

Qy 1 YGRKKRQRRR 11
| | | | |
Db 47 YGRKKRQRRR 57

RESULT 11

S54381

tat protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S54381
R:Theodore, T.; Buckler-White, A.J.
submitted to the EMBL Data Library, July 1989
A:Reference number: S54377
A:Accession: S54381
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-86 <THE>
A:Cross-references: UNIPROT:P12506; EMBL:M22639; NID:G329377; PIDN:AAA45363.1; PID:G32937
C:Genetics:
A:Introns: 72/2
C:Superfamily: AIDS trans-activating transcription regulator

Query Match 37.9%; Score 58; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.74; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

Qy 1 YGRKKRQRRR 11

db 47 YGRKRRRR 57

RESULT 12

RESULT 12
S33982
trans-activating transcription regulator - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:date: 08-Oct-1994 #sequence revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: S33982; S26385; S19864

R:Carlini, F.

submitted to the EMBL Data Library, November 1991

A;Reference number: S33979
A;Accession: S33982
A;Molecule type: DNA
A;Residues: 1-86 <AR>
A;Cross-references: UNIPROT:P04606; EMBL:Z11530; NID:g60192; PID:g60196
R;Sliderovski, D.P.; Matsuyama, T.; Frigerio, E.; Chui, S.; Min, X.; Efile, H.; Sumner-Smith, J.; Brown, R.E.; Bressan, G.; Bressan, M.; Bressan, P.; Bressan, R.; Bressan, S.;
Nucleic Acids Res. 20, 5311-5320, 1992
A;Title: Random mutagenesis of the human immunodeficiency virus type-1 trans-activator protein (Tat) gene
A;Reference number: S26385; MUID:93065196; PMID:1437550
A;Accession: S26385
A;Molecule type: nucleic acid
A;Residues: 1-86 <SID>
A;Cross-references: EMBL:X64650; NID:g60144; PID:CAA45921.1; PID:g60145

C;Genetics:
A;Gene: tat
A;Introns: 72/2
C;Superfamily: AIDS trans-activating transcription regulator
C;Keywords: AIDS; immunodeficiency

Query Match	37.9%	Score 58;	DB 2;	Length 86;
Best Local Similarity	100.0%;	Pred. No. 0.74;		
Matches 11: Conservative	0;	Mismatches	0;	Indels 0;
Gaps	0;			

QY 1 YGRKKRRQRRR 11
Db 47 YGRKKRRQRRR 57

RESULT 13

T01665
tat protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01665
R:Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.
Cell 46, 63-74, 1986
A:Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isolates
A:Reference number: Z14389; MUID:86245056; PMID:2424612
A:Accession: T01665
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-87 <ALI>
A:Cross-references: UNIPROT:P04613; EMBL:K03456; NID:g60228; PID:CAA28015.1; PID:g60233
C:Genetics:
A:Introns: 72/2
C:Superfamily: AIDS trans-activating transcription regulator

Query Match 37.9%: Score 58; DB 2: Length 87;

Best Local Similarity	100.0%;	Pred. No. 0.75;	
Matches	11;	Conservative	0;
Mismatches	0;	Mismatches	0;
Indels	0;	Indels	0;
Gaps	0;	Gaps	0;

Qy 1 YGRKRRQRRR 11
Db 47 YGRKRRQRRR 57

RESULT 14

trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate TNL012)
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)

C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004
C;Accession: A04017
R;Arya, S.K.; Gallo, R.C.
Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986
A;Title: Three novel genes of human T-lymphotropic virus type III: immune r
A;Reference number: A94093; MUID:86177573; PMID:3008154

A;Accession: A04017
A;Molecule type: DNA
A;Residues: 1-95 <ARY>
A;Cross-references: UNIPROT: P04326

C; Disease: AIDS
C; Genetic: tat
A; Gene: tat
C; Superfamily: AIDS trans-activating transcription regulator
C; Keywords: AIDS; immunodeficiency; transcription regulation

Query Match	37.9%	Score 58;	DB 1;	Length 95;
Best Local Similarity	100.0%;	Pred. No. 0.82;		
Matches 11: Conservative	0: Mismatches	0: Indels	0: Gaps	0: Gaps

Qy 1 YGRKRRQRRR 11
p'b 56 YGRKRRQRRR 66

RESULT 15

trans-activating transcription regulator - human immunodeficiency virus type 1 (strain Y83001)
N/Alternate names: tat protein
C/Species: human immunodeficiency virus type 1, HIV-1
A/Note: host Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: E44001
R/Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M. J. Virol. 66, 6587-6600, 1992
A/Title: Complete nucleotide sequence, genome organization, and biological properties of A/Reference number: A44001; MUID:93021387; PMID:1404605

A;Accession: E44001
A;Molecule type: DNA
A;Residues: 1-101 <LTY>
A;Cross-references: UNIPROT:P15965; GB:M93258
C;Genetics:
A;Gene: tat

Query Match	37.9%	Score 58;	DB 1;	Length 101;
Best Local Similarity	100.0%;	Pred. No. 0.88;		
Matches 11: Conservative	0;	Mismatches	0;	Indels 0;
Matches 11: Conservative	0;	Mismatches	0;	Indels 0;
Matches 11: Conservative	0;	Mismatches	0;	Indels 0;

Qy 1 YGRKRRQRRR 11
D'b 47 YGRKRRQRRR 57

RESIST 16

tat protein - human immunodeficiency virus type 1 (strain JRFL)
 C/Species: human immunodeficiency virus type 1, HIV-1
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C/Accession: T09446
 R/Pang, S.; Winters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, I.
 Submitted to the EMBL Data Library. July 1996

A:Accession: T09446
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-101 <PAN>
 A:Reference number: Z16673
 Submitted to the EMBL data library, July 1990

A; Coordinates: UNIPROT:Q75758; EMBL:U63632; NID:gl46577
C; Geneticals:
A; Gene: tat
C; Superfamily: AIDS trans-activating transcription regulator
A; Introns: 72/2

```

hypothetical protein BH0732 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: D83741
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira-
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D83741
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-457 <STO>
A;Cross-references: UNIPROT:Q9KEW7; GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BA804.3
A;Experimental source: strain C-125
C:Genetics:
A;Gene: BH0732

Query Match 34.0%; Score 52; DB 2; Length 457;
Best Local Similarity 40.0%; Pred. No. 25;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 GRKKRRRRRLDLEMLAXYIPMDD 26
      ||| ||| :| :| :| |||
Db 202 GKVKRKRTGSELHAIRYIPDD 226

RESULT 20
E70610
hypothetical protein Rv1215c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70610
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70610
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-561 <CLO>
A;Cross-references: UNIPROT:O05316; GB:Z93777; GB:AL123456; NID:g3261726; PIDN:CAB07817
A;Experimental source: strain H37RV
C:Genetics:
A;Gene: Rv1215c

Query Match 34.0%; Score 52; DB 2; Length 561;
Best Local Similarity 35.7%; Pred. No. 32;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 YGRKKRRQRRRLDLEMLAXYIPMDDDF 28
      :| ||| :| :| :| :|
Db 229 WSGVRRTTRQTYDMEQOEAHPLRDDF 256

RESULT 21
T01773
syntaxin homolog A.IG002P16.16 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 24-Mar-1999
C:Accession: T01773
R;Miller, N.; Beck, C.; Kramer, J.
submitted to the EMBL Data Library, June 1997
A;Description: The sequence of A. thaliana IG002P16.
A;Reference number: Z14421
A:Accession: T01773
A;Status: translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-307 <MIL>
A;Cross-references: EMBL:AF007270; NID:g2191157; PID:g2191179
A;Experimental source: cultivar Columbia

```

C;Genetics:
A;Map position: 5
A;Introns: 67/2; 159/3; 189/3; 213/3; 249/3; 271/3
A;Note: A_IG002P16.16

Query Match 33.7%; Score 51.5; DB 2; Length 307;
Best Local Similarity 38.7%; Pred. No. 20;
Matches 12; Conservative 7; Mismatches 9; Indels 3; Gaps 1;

QY 1 YGRKKRRQRRLDLEW--LAXYIPMDDDF 28
| : : | : : ||||| : : ||||
DB 180 YLKRLRQQKEDGMDLENLSNRYRPEDDF 210

RESULT 22
F64651
hypothetical protein HP1054 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: F64651
R;Tomb, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: F64651
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-422 <TOM>
A;Cross-references: UNIPROT:O25694; GB:A5000613; GB:A5000511; NID:g2314200; PIDN:AAD0810
C;Genetics:
A;Start codon: TTG

```

Query Match          33.3%; Score 51; DB 2; Length 422;
Best Local Similarity 37.9%; Pred. No. 33;
Matches 11; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy      1 YGRKKRRQRRDLLEMLAXYIPMDDDFQ 29
       :|||:|||:|||||:|
Db     215 FKRKIHLREKIDLSALKOKIAQKEKFQ 243

RESULT 23
G71939
hypothetical protein jhp0371 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: G71939
R:AJM, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho-
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: G71939
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-440 <ARN>
A:Cross-references: UNIPROT:Q9ZM52; GB:AE001472; GB:AE001439; NID:G4154897; PIDN:AAD0595
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0371

```

Query Match 33.3%; Score 51; DB 2; Length 440;
Best Local Similarity 37.9%; Pred.No. 34;
Matches 11; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

```
QY      1 YGRKRRQRRLDLEMLAXYPMDDDFQ 29  
        :|::||::||::||  
DB     233 FKRTHRLRERDIINLSAKDKIAQEKFQ 261
```

RESULT 24
 C85651
 hypothetical protein Z1556 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: C85651
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: C85651
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-462 <STO>
 A:Cross-references: UNIPROT:C8XAK4; GB:AE005174; NID:G12514427; PIDN:AAG55671.1; GSPDB:GN
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z1556

Query Match	33.0%	Score 50.5;	DB 2;	Length 462;
Best Local Similarity	37.9%;	Pred. No. 42;		
Matches 11;	Conservative 6;	Mismatches 11;	Indels 1;	Gaps 1;
Qy	1 YGRKKRRQRRLD-LEMLAXYIPMDDDF	28		
	: : : :	:	:	:
Db	330 YSRAEQRORELIDFLNTTGYAPLDQAF	358		

RESULT 25
H90790
hypothetical protein ECs1296 [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90790
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic
A:Reference number: A39629; MUID:21156231; PMID:11258796
A:Accession: H90790
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <HAY>
A:Cross-references: UNIPROT:Q8XAK4; GB:BA000007; PIDN:BA334719.1; PID:gl3360756; GSPDB:G
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs1296

	Query Match	33.0%;	Score 50.5;	DB 2;	Length 462;	
	Best Local Similarity	37.9%;	Pred. No. 42;			
	Matches 11;	Conservative	6;	Mismatches 11;	Indels 1;	Gaps 1;
Qy	1 YGRKKRRQRRRLD-LEMLAXYIPMDDDF	28				
	: : : : :					
Dp	330 YSRAQRQORELIDFLNTTTGYAPLDQAF	358				

RESULT 26
A70966
hypothetical protein Rv2653c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: A70966
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70966

A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-107 <COL>
A:Cross-references: UNIPROT:P71950; GB:Z80225; GB:AL123456; NID:G3242265; PIDN:CA802358
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV253c
C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV2653c

Query Match 32.7%; Score 50; DB 2; Length 107;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 12; Conservative 7; Mismatches 5; Indels 6; Gaps 1;

QY 4 KKKRRRRRLD-----LEMLAXYIPMDD 27
|||:||||: |||: |||:
Db 39 QRRQRQDLEAIRRAYAEWVATSHIEDD 68

RESULT 27
RWUC2
T-cell surface glycoprotein CD2 precursor - human
N:Alternate names: E rosette receptor; erythrocyte receptor; erythrocyte-binding protein
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1989 #sequence revision 31-Mar-1989 #text change 09-Jul-2004
C:Accession: A28967; B26486; A28416; A28023; S02292; A30430; S00829; A29874
R:Diamond, D.J.; Clayton, L.K.; Sayre, P.H.; Reinherz, E.L.
Proc. Natl. Acad. Sci. U.S.A. 85, 1615-1619, 1988
A:Title: Exon-intron organization and sequence comparison of human and murine T11 (CD2)
A:Reference number: A28967; MUID:88144486; PMID:2894031
A:Accession: A28967
A:Molecule type: DNA
A:Residues: 1-351 <DIA>
A:Cross-references: UNIPROT:P06729; GB:M19806; GB:J03622; GB:J03623; NID:G180079; PIDN:A
R:Sewell, W.A.; Brown, M.H.; Dunne, J.; Owen, M.J.; Crumpton, M.J.
Proc. Natl. Acad. Sci. U.S.A. 83, 8718-8722, 1986
A:Title: Molecular cloning of the human T-lymphocyte surface CD2 (T11) antigen.
A:Reference number: A26486; MUID:87041523; PMID:3490670
A:Accession: A26486
A:Molecule type: mRNA
A:Residues: 1-338, 'M', 340, 'Q'QKTHCPLPLIKKDRNCLFQ' <SE1>
A:Accession: B26486
A:Molecule type: protein
A:Residues: 25-46, 'X', 50 <SE2>
R:Sewell, W.A.; Brown, M.H.; Dunne, J.; Owen, M.J.; Crumpton, M.J.
Proc. Natl. Acad. Sci. U.S.A. 84, 7256, 1987
A:Reference number: A28416
A:Contents: revision
A:Accession: A28416
A:Molecule type: mRNA
A:Residues: 333-351 <SE3>
R:Seed, B.; Aruffo, A.
Proc. Natl. Acad. Sci. U.S.A. 84, 3365-3369, 1987
A:Title: Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rap
A:Reference number: A28023; MUID:87204137; PMID:2437578
A:Accession: A28023
A:Molecule type: mRNA
A:Residues: 1-265, 'Q', 267-351 <SEE>
A:Cross-references: GB:M16445; NID:G178668; PIDN:AAA51738.1; PID:G178669
R:Sayre, P.H.; Chang, H.C.; Hussey, R.E.; Brown, N.R.; Richardson, N.E.; Spagnoli, G.; C
Proc. Natl. Acad. Sci. U.S.A. 84, 2941-2945, 1987
A:Title: Molecular cloning and expression of T11 cDNAs reveal a receptor-like structure
A:Reference number: S02292; MUID:87024243; PMID:2883656
A:Accession: S02292
A:Molecule type: mRNA
A:Residues: 1-338, 'M', 340, 'Q'QKTHCPLPLIKKDRNCLFQ' <SA1>
A:Cross-references: GB:M16336; NID:G180093; PIDN:AAA51946.1; PID:G180094
A:Accession: A30430
A:Molecule type: protein
A:Residues: 25-43, 152-163 <SA2>
R:Lang, G.; Wotton, D.; Owen, M.J.; Sewell, W.A.; Brown, M.H.; Mason, D.Y.; Crumpton, M.
EMBO J. 7, 1675-1682, 1988
A:Title: The structure of the human CD2 gene and its expression in transgenic mice.
A:Reference number: S00829; MUID:89005055; PMID:2901953

A:Accession: S00829
A:Molecule type: DNA
A:Residues: 1-351 <LAN>
A:Cross-references: EMBL:X07871
C:Comment: CD2 is a surface antigen expressed on all peripheral blood T-cells. It appear
or is closely associated with, the erythrocyte receptor.
C:Genetics:
A:Gene: GDB:CD2
A:Cross-references: GDB:I18735; OMIM:186990
A:Map position: lp13.1-1p13.1
A:Introns: 21/1; 128/1; 205/1; 246/1
C:Superfamily: T-cell surface glycoprotein CD2
C:Keywords: glycoprotein; T-cell; transmembrane protein
P:1-24/Domain: signal sequence #status predicted <SIG>
P:25-351/Product: T-cell surface glycoprotein CD2 #status predicted <MAT>
P:25-206/Domain: extracellular #status predicted <EXT>
P:210-234/Domain: transmembrane #status predicted <TMM>
P:237-351/Domain: intracellular #status predicted <INT>
P:89,141,150/Binding site: carbohydrate (Asn) #status predicted

Query Match 32.7%; Score 50; DB 1; Length 351;
Best Local Similarity 41.7%; Pred. No. 37;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 3 KKKRRRRRLDLEMLAXYIPMDD 26
|||:||||: |||: |||:
Db 237 KKKRRRRRLDLEMLAXYIPMDD 260

RESULT 28
S35637
high mobility group 1 protein homolog - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text change 16-Aug-2004
C:Accession: S35637; S35636; I84754
R:Wang, L.; Precht, P.; Balakir, R.; Horton Jr., W.E.
submitted to the EMBL Data Library, January 1993
A:Reference number: S35637
A:Accession: S35637
A:Molecule type: mRNA
A:Residues: 1-561 <WAN>
A:Cross-references: UNIPROT:Q04931; GB:L08814; NID:G203464; PIDN:AAA40927.1; PID:G203465
R:Wang, L.; Precht, P.; Balakir, R.; Horton Jr., W.E.
Nucleic Acids Res. 21, 1493, 1993
A:Title: Rat and chick cDNA clones encoding HMG-like proteins.
A:Reference number: I50198; MUID:93219134; PMID:8464746
A:Accession: S35636
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 391-476 <MAW>
A:Cross-references: GB:L08814; NID:G203464; PIDN:AAA40927.1; PID:G203465
C:Genetics:
A:Gene: C11DBP
C:Superfamily: HMG box homology
C:Keywords: DNA binding
P:396-471/Domain: HMG box homology <HMG1>

Query Match 32.7%; Score 50; DB 2; Length 561;
Best Local Similarity 45.8%; Pred. No. 60;
Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 4 KKKRRRRRLDLEMLAXYIPMDD 27
|||:||||: |||: |||:
Db 531 KKKRRRRRLDLEMLAXYIPMDD 554

RESULT 29
A28199
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha chain - chicken
C:Species: Gallus gallus (chicken)
C:Date: 21-Sep-1988 #sequence revision 21-Sep-1988 #text change 09-Jul-2004
A:Accession: A28199
R:Takeyasu, K.; Tamkun, M.M.; Renaud, K.J.; Fambrough, D.M.

J. Biol. Chem. 263, 4347-4354, 1988
A>Title: Ouabain-sensitive (Na⁺) + K⁺) -ATPase activity expressed in mouse L cells by
A;Reference number: A28199; MUID:89153759; PMID:2831227
A;Accession: A28199
A>Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1021 <TAK>
A;Cross-references: UNIPROT:P09572; GB:J03230; NID:G211219; PIDN:AAA48607.1; PID:G211220
C;Superfamily: Na⁺/K⁺-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: ATP; glycoprotein; hydrolase; phosphoprotein; transmembrane protein
F;585-781/Domain: ATPase nucleotide-binding domain homology <ATN>
F;213, 481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;374/Active site: Asp (aspartylphosphate intermediate) #status predicted
F;506/Binding site: ATP (lys) #status predicted

Query Match 32.7%; Score 50; DB 2; Length 1021;
Best Local Similarity 42.3%; Pred. No. 1.1e+02;
Matches 11; Conservative 6; Mismatches 7; Indels 2; Gaps 1;
Qy 1 YGRKKRRQRRDLLEMLAXYIPMD 26
Db 18 HGTKKKAKERDMD--ELKKSIMDD 41

RESULT 30
A24639
Na⁺/K⁺-exchanging ATPase (EC 3.6.3.9) alpha-1 chain [validated] - rat
N/Alternate names: Na⁺/K⁺-transporting ATPase alpha chain, kidney-type
C;Contains: Na⁺/K⁺-transporting ATPase alpha-S chain
N;Species: Rattus norvegicus (Norway rat)
C;Date: 18-Aug-2000 #sequence revision 18-Aug-2000 #text change 09-Jul-2004
C;Accession: A24639; S00460; A2180; S11020; A25171; S29877; S10758
R;Shull, G.E.; Greeb, J.; Lingrel, J.B.
Biochemistry 25, 8125-8132, 1986
A>Title: Molecular cloning of three distinct forms of the Na⁺, K⁺-ATPase alpha-subunit fr
A;Reference number: A30512; MUID:87128908; PMID:3028470
A;Accession: A24639
A;Molecule type: mRNA
A;Residues: 1-1023 <SHU>
A;Cross-references: UNIPROT:P06685; EMBL:M14511; NID:G203026; PIDN:AAA40775.1; PID:G2030
R;Hata, Y.; Urayama, O.; Kawakami, K.; Nojima, H.; Nagamune, H.; Kojima, T.; Ohta, T.; N
J. Biochem. 102, 43-58, 1987
A>Title: Primary structures of two types of alpha-subunit of rat brain Na⁺(+), K⁺(+)-ATPase
A;Reference number: S00460; MUID:88032933; PMID:2822682
A;Accession: S00460
A;Molecule type: mRNA
A;Residues: 1-1023 <HAR>
A;Cross-references: EMBL:X05882; NID:G55771; PIDN:CAA29306.1; PID:G55772
R;Herrera, V.L.M.; Emanuel, J.R.; Ruiz-Opazo, N.; Levenson, R.; Nadal-Ginard, B.
J. Cell Biol. 105, 1855-1865, 1987
A>Title: Three differentially expressed Na, K-ATPase alpha subunit isoforms: structural a
A;Reference number: A92749; MUID:88033255; PMID:2822726
A;Accession: A27180
A;Molecule type: mRNA
A;Residues: 1-67, 'pv', 70-174, 'E', 176-187, 'V', 189-334, 'V', 336-1023 <HER>
A;Cross-references: EMBL:M28647; NID:G205631; PIDN:AAA41671.1; PID:G205632
R;Yagawa, Y.; Kawakami, K.; Nagano, K.
Biochim. Biophys. Acta 1049, 286-292, 1990
A>Title: Cloning and analysis of the 5'-flanking region of rat Na⁺(+), K⁺(+)-ATPase alpha-1
A;Reference number: S11020; MUID:90344872; PMID:2166579
A;Accession: S11020
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-41 <YAG>
A;Cross-references: EMBL:X53233
R;Schneider, J.W.; Mercer, R.W.; Caplan, M.; Emanuel, J.R.; Sweadner, K.J.; Benz Jr., E.
Proc. Natl. Acad. Sci. U.S.A. 82, 6357-6361, 1985
A>Title: Molecular cloning of rat brain Na, K-ATPase alpha-subunit cDNA.
A;Reference number: A25171, MUID:85298352; PMID:2994074
A;Accession: A25171
A;Molecule type: mRNA
A;Residues: 489-533 <SCH>
R;Lytton, J.

Biochem. Biophys. Res. Commun. 132, 764-769, 1985
A>Title: The catalytic subunits of the (Na⁺), K⁺(+)-ATPase alpha and alpha(+) isozymes
A;Reference number: S29877; MUID:86050667; PMID:2998384
A;Accession: S29877
A>Status: preliminary
A;Molecule type: protein
A;Residues: 6-19 <LYT>
R;Kurihara, K.; Hosoi, K.; Kodama, A.; Ueha, T.
Biochim. Biophys. Acta 1039, 234-240, 1990
A>Title: A new electrophoretic variant of alpha subunit of Na⁺(+), K⁺(+)-ATPase from the su
A;Reference number: S10758; MUID:90304196; PMID:2163680
A;Accession: S10758
A;Molecule type: protein
A;Residues: 6, 'X', 8-10, 'X', 12-16 <KUR>
A;Experimental source: submandibular gland
A;Note: designated alpha-S form; thought to arise from alpha-1 chain by post-translation
C;Genetics:
C;Gene: NKAAL
A;Introns: 4/3
A;Note: the list of introns may be incomplete
C;Superfamily: Na⁺/K⁺-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp
F;6-1023/Product: Na⁺/K⁺-transporting ATPase alpha-1 chain #status experimental <MAT>
F;6-95/Domain: intracellular #status predicted <INT1>
F;96-120/Domain: transmembrane #status predicted <TM1>
F;130-149/Domain: transmembrane #status predicted <TM2>
F;150-290/Domain: intracellular #status predicted <INT2>
F;231-313/Domain: transmembrane #status predicted <TM3>
F;340-348/Domain: transmembrane #status predicted <TM4>
F;349-786/Domain: intracellular #status predicted <INT3>
F;587-783/Domain: ATPase nucleotide-binding domain homology <ATN>
F;787-810/Domain: transmembrane #status predicted <TM5>
F;849-874/Domain: transmembrane #status predicted <TM6>
F;875-952/Domain: intracellular #status predicted <INT4>
F;953-978/Domain: transmembrane #status predicted <TM7>
F;979-1023/Domain: extracellular #status predicted <EXT>
F;376/Active site: Asp (aspartylphosphate intermediate) #status predicted
F;508/Binding site: ATP (Lys) #status predicted
F;717, 721, 726/Active site: Asp, Asp, Lys #status predicted

Query Match 32.7%; Score 50; DB 1; Length 1023;
Best Local Similarity 30.8%; Pred. No. 1.1e+02;
Matches 8; Conservative 10; Mismatches 8; Indels 0; Gaps 0;
Qy 1 YGRKKRRQRRDLLEMLAXYIPMD 26
Db 18 HGDKKKAKKERDMDDELKKEVSMD 43

RESULT 31
A24414
Na⁺/K⁺-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - human
N/Alternate names: sodium pump; sodium/potassium transporting ATPase alpha-A chain
C;Species: Homo sapiens (man)
C;Date: 02-Jun-1988 #sequence revision 02-Jun-1988 #text_change 09-Jul-2004
C;Accession: A24414; A27795; A39910; I60116; S09171
R;Kawakami, K.; Ohta, T.; Nojima, H.; Nagano, K.
J. Biochem. 100, 389-397, 1986
A>Title: Primary structure of the alpha-subunit of human Na, K-ATPase deduced from cDNA s
A;Reference number: A24414; MUID:87057096; PMID:2430951
A;Accession: A24414
A;Molecule type: mRNA
A;Residues: 1-1023 <KAW>
A;Cross-references: UNIPROT:P05023; EMBL:X04297; NID:G28926; PIDN:CAA27840.1; PID:G28927
R;Shull, M.M.; Lingrel, J.B.
Proc. Natl. Acad. Sci. U.S.A. 84, 4039-4043, 1987
A>Title: Multiple genes encode the human Na⁺, K⁺-ATPase catalytic subunit.
A;Reference number: A94158; MUID:87231946; PMID:3035563
A;Accession: A27795
A;Molecule type: DNA
A;Residues: 168-189; 213-214, 'X', 216-244 <SHU>
R;Chehab, F.F.; Kan, Y.W.; Law, M.L.; Hartz, J.; Kao, F.T.; Blostein, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 7901-7905, 1987

Query Match	32.4%	Score 49.5;	DB 2;	Length 406;
Best Local Similarity	47.8%	Pred. No. 51;		
Matches 11;	Conservative	4;	Mismatches 7;	Indels 1;
Matches 1;	Gaps			

C; Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
R; Accession: B82444
R; Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson
R; Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson

1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: B82444
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <HEI>
A:Cross-references: UNIPROT:Q9KM28; GB:AE004387; GB:AE003853; NID:g9657967; PIDN:AAF9646
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0561
A:Map position: 2

Query Match 32.0%; Score 49; DB 2; Length 124;
Best Local Similarity 37.5%; Pred. No. 18;
Matches 9; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 7 RRRRDLLEMLAXYIPMDDDFOL 30
DB 86 RRRRDSVPVIFTHARFDDDAEL 109

RESULT 36
D90478
conserved hypothetical protein [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
A:Accession: D90478
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: D90478
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <KUR>
A:Cross-references: UNIPROT:Q9TUM5; GB:AE006641; NID:g13816361; PIDN:AAK43083.1; GSPDB:C
C:Genetics:
A:Gene: SSO2979

Query Match 32.0%; Score 49; DB 2; Length 320;
Best Local Similarity 32.4%; Pred. No. 46;
Matches 11; Conservative 9; Mismatches 10; Indels 4; Gaps 1;

QY 1 YGRKKRRQRRDL-----LEMLAXYIPMDDDFOL 30
DB 188 YNKKENKQHRDLNTLVVVGMSAFYSPNNKIEL 221

RESULT 37
T03776
tat binding protein homolog - rice
C:Species: *Oryza sativa* (rice)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
A:Accession: T03776
R:Suzuka, I.; Koga-Ban, Y.; Sasaki, T.; Minobe, Y.; Hashimoto, J.
Plant Sci. 103, 33-40, 1994
A:Title: Identification of cDNA clones for rice homologs of the human immunodeficiency v
A:Reference number: Z15075
A:Accession: T03776
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-448 <SUZ>
A:Cross-references: UNIPROT:P46466; EMBL:D17789; PIDN:BAA04615.1
A:Experimental source: cv. Nipponbare, callus
C:Superfamily: ATP-dependent 26S proteinase; Ftsh/SEC18/CDC48-type ATP-binding domain h
F:207-417/Domain: Ftsh/SEC18/CDC48-type ATP-binding domain homology <VAT>

Query Match 32.0%; Score 49; DB 2; Length 448;
Best Local Similarity 28.9%; Pred. No. 65;

Matches 13; Conservative 8; Mismatches 6; Indels 18; Gaps 1;

QY 2 GRKKRQ-----RRDLLEMLAXYIPMDDDF 28
DB 42 GRKQKQKQGEAAARLPNAPVLSKRLRLKLERVKDYLLMEEF 86

RESULT 38
S04630
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-1 chain - horse
C:Species: *Equus caballus* (domestic horse)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
A:Accession: S04630
R:Kano, I.; Nagai, F.; Satoh, K.; Ushiyama, K.; Nakao, T.; Kano, K.
FEBS Lett. 250, 91-98, 1989
A:Title: Structure of the alpha(1) subunit of horse Na,K-ATPase gene.
A:Reference number: S04630; MUID:89290042; PMID:2544461
A:Accession: S04630
A:Molecule type: DNA
A:Residues: 1-1021 <KAN>
A:Cross-references: UNIPROT:P18907; EMBL:X16773; NID:g1010; PIDN:CAA34716.1; PID:g871026
C:Genetics:
A:Introns: 4/3; 39/3; 59/3; 127/3; 165/3; 210/3; 250/3; 339/3; 406/1; 442/3; 487/3; 552/
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp
F:6-1021/Product: Na+/K+-transporting ATPase alpha-1 chain #status predicted <MAT>
F:6-93/Domain: intracellular #status predicted <INT1>
F:94-118/Domain: transmembrane #status predicted <TM1>
F:118-147/Domain: transmembrane #status predicted <TM2>
F:148-288/Domain: intracellular #status predicted <INT2>
F:289-311/Domain: transmembrane #status predicted <TM3>
F:318-346/Domain: transmembrane #status predicted <TM4>
F:347-784/Domain: intracellular #status predicted <INT3>
F:585-781/Domain: ATPase nucleotide-binding domain homology <ATN>
F:785-808/Domain: transmembrane #status predicted <TM5>
F:847-872/Domain: transmembrane #status predicted <TM6>
F:873-950/Domain: intracellular #status predicted <INT4>
F:951-976/Domain: transmembrane #status predicted <TM7>
F:977-1021/Domain: extracellular #status predicted <EXT>
F:374/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:506/Binding site: ATP (Lys) #status predicted
F:715,719,724/Active site: Asp, Asp, Lys #status predicted

Query Match 32.0%; Score 49; DB 1; Length 1021;
Best Local Similarity 38.5%; Pred. No. 1.5e+02;
Matches 10; Conservative 7; Mismatches 7; Indels 2; Gaps 1;

QY 1 YGRKKRRQRRDLLEMLAXYIPMD 26
DB 18 HGNKKKAKGRDMD--ELKKEVSMDD 41

RESULT 39
S26180
neurofascin - chicken
C:Species: *Gallus gallus* (chicken)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
A:Accession: S26180
R:Volkmer, H.; Hassel, B.; Wolff, J.M.; Frank, R.; Rathjen, F.G.
J. Cell Biol. 118, 149-161, 1992
A:Title: Structure of the axonal surface recognition molecule neurofascin and its relati
A:Reference number: S26180; MUID:92317154; PMID:1377696
A:Accession: S26180
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1272 <VOL>
A:Cross-references: UNIPROT:Q42414; EMBL:X65224; NID:g63659; PIDN:CAA46330.1; PID:g63660
C:Superfamily: neural cell adhesion molecule I1; fibronectin type III repeat homology; I
F:279-336/Domain: immunoglobulin homology <INM>

Query Match 32.0%; Score 49; DB 2; Length 1272;
Best Local Similarity 43.3%; Pred. No. 1.9e+02;
Matches 13; Conservative 5; Mismatches 6; Indels 6; Gaps 1;

A;Accession: S74207
A;Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A;Residues: 1-878 <HOR>
A;Cross-references: UNIPROT:Q42710; EMBL:X92890; NID:g1296511; PIDN:CAA63483.1
A;Experimental source: tissue cotedones; clone PCSLBLOX221
A;Accession: S741137
A:Molecule type: protein
A;Residues: 196,'X',198-204;218,'FX',221-223,'XX',226-228;305-309,'XX',312-316;441-460;8;
C;Superfamily: lipoxigenase
C;Keywords: oxidoreductase

Query Match 31.7%; Score 48.5; DB 2; Length 878;
Best Local Similarity 35.3%; Pred. NO. 1.5e+02;
Matches 12; Conservative 6; Mismatches 9; Indels 7; Gaps 1;

Qy 2 GRKRRQRRLDLLE-----MLAXYIPMDDDF 28
||| : ||| : ||| : ||| : ||| :
Db 251 GRTGPRSRDHNYESRLSPIMSLDIYVPEKNF 284

RESULT 43
AH3337
transcription regulatory protein, lysr family BMEI0686 [imported] - Brucella melitensis
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AH3337
R;DeVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AH3337
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-316 <KUR>
A;Cross-references: UNIPROT:O8YHW0; GB:AE008917; PIDN:AAL51867.1; PID:g17982617; GSPDB:
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0686
A;Map position: I
C;Superfamily: hypothetical protein b1875

Query Match 31.4%; Score 48; DB 2; Length 316;
Best Local Similarity 32.1%; Pred. NO. 63;
Matches 9; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Qy 3 RKRRQRRLDLEMLAXYIPMDDDFL 30
||| : ||| : ||| : ||| : ||| :
Db 152 RKLLDERLDVTLVNISQAVPTRDDVEI 179

RESULT 44
I49585
CD2 antigen precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49585
R;Yagita, H.; Okumura, K.; Nakauchi, H.
J. Immunol. 140, 1321-1326, 1988
A;Title: Molecular cloning of the murine homologue of CD2: Homology of the molecule to 1
A;Reference number: I49585; MUID:88140313; PMID:3257775
A;Accession: I49585
A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 1-344 <RES>
A;Cross-references: UNIPROT:P08920; GB:M18934; NID:g192486; PIDN:AAA37397.1; PID:g309158
C;Superfamily: T-cell surface glycoprotein CD2

Query Match 31.4%; Score 48; DB 2; Length 344;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 RKRRQRRLDLEMLAX 20
|:|:|:|:|:|:|:
Db 231 RKRRNRKDELEIKAS 248

RESULT 45
B28967
T-cell surface glycoprotein CD2 precursor - mouse
N:Alternate names: CD2 antigen; T-lymphocyte antigen CD2; T11 protein
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1989 #sequence_revision 03-Jun-1993 #text_change 09-Jul-2004
C:Accession: B28967; S01347; S02293
R:Diamond, D.J.; Clayton, L.K.; Sayre, P.H.; Reinherz, E.L.
Proc. Natl. Acad. Sci. U.S.A. 85, 1615-1619, 1988
A:Title: Exon-intron organization and sequence comparison of human and murine T11 (CD2)
A:Reference number: A28967; MUID:98144486; PMID:2894031
A:Accession: B28967
A:Molecule type: mRNA
A:Residues: 1-344 <DTA>
A:Cross-references: UNIPROT:Q9R201; GB:M19807; NID:G192479; PIDN:AAA37393.1; PID:G3871222
A:Note: The authors translated the codon TAT for residue 99 as Thr
R:Clayton, L.K.; Sayre, P.H.; Novotny, J.; Reinherz, E.L.
Eur. J. Immunol. 17, 1367-1370, 1987
A:Title: Murine and human T11 (CD2) cDNA sequences suggest a common signal transduction
A:Reference number: S01347; MUID:98004738; PMID:2820751
A:Accession: S01347
A:Molecule type: mRNA
A:Residues: 1-127,'M',129-174,'N',176-190,'NW',193-344 <CLA>
A:Cross-references: EMBL:X06143; NID:G54223; PIDN:CAA9500.1; PID:G54224
R:Sewell, W.A.; Brown, M.H.; Owen, M.J.; Fink, P.J.; Kozak, C.A.; Crumpton, M.J.
Eur. J. Immunol. 17, 1015-1020, 1987
A:Title: The murine homologue of the T lymphocyte CD2 antigen: molecular cloning, chromo-

Query Match 31.4%; Score 48; DB 2; Length 344;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 RKRRQRRLDLEMLAX 20
|:|:|:|:|:|:|:
Db 231 RKRRNRKDELEIKAS 248

RESULT 46
S58152
Hypothetical protein SPAC2F7.08c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: T38556; S58152
R:Gentiles, S.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z21799
A:Accession: T38556
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-632 <GE2>
A:Cross-references: UNIPROT:Q09699; EMBL:Z50142; NID:G1052783; PIDN:CAA90495.1; PID:G1052783
A:Experimental source: strain 972h-; cosmid c2F7
C:Genetics:

Job time : 33.0526 secs

```
RESULT 49
H86105
probable vimentin yJda [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H86105
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H86105
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <STO>
A:Cross-references: UNIPROT:O8XDV5; GB:AE005174; NID:G12519079; PIDN:AGS59308.1; GSPDB:C
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yJda
C:Superfamily: Escherichia coli hypothetical protein b4109

Query Match 31.4%; Score 48; DB 2; Length 742;
Best Local Similarity 34.6%; Pred. No. 1.5e+02;
Matches 9; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 4 KRRQRRLDLEMLAXYIPMDDDFQ 29
Db 133 QKLQRLRDCDIKHLTDVLEIDKMR 158

RESULT 50
S29344
protein kinase KIN3 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O5220; protein YOR233w
C:Species: Saccharomyces cerevisiae
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-2004
A:Accession: S29344; S67126; S24707
R:Kambouris, N.G.; Burke, D.J.; Creutz, C.E.
Yeast 9, 141-150, 1993
A:Title: Cloning and genetic analysis of the gene encoding a new protein kinase in Sacch
A:Reference number: S29344; MUID:93220392; PMID:8465601
A:Accession: S29344
A:Molecule type: DNA
A:Residues: 1-800 <KAM>
A:Cross-references: UNIPROT:O01919; EMBL:X67916; NID:G5514; PIDN:CAA48115.1; PID:G5515
R:Boyer, J.; Fairhead, C.; Gaillon, L.; Galisson, F.; Michaux, G.; Thierry, A.; Dujon, E
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67104
A:Accession: S67126
A:Molecule type: DNA
A:Residues: 1-800 <BOY>
A:Cross-references: EMBL:Z75141; NID:G1420534; PIDN:CAA99453.1; PID:e252094; PID:G142053
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:KIN4; KIN3; KIN31
A:Cross-references: SGD:S0005759; MIPS:YOR233w
A:Map position: 15R
C:Superfamily: protein kinase homology
K:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:44-313/Domain: protein kinase homology <KIN>
F:52-60/Region: protein kinase ATP-binding motif

Query Match 31.4%; Score 48; DB 2; Length 800;
Best Local Similarity 61.5%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 17 MLAXYIPMDDDFQ 29
Db 245 MLAGYLPWDDHE 257
```

Search completed: February 8, 2005, 20:32:30

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 05:55:37 ; Search time 31.0526 Seconds
(without alignments)
72.119 Million cell updates/sec

Title: US-10-032-361-7
Perfect score: 159
Sequence: 1 YGRKRRQRRLDLEMLAPYPMDDFQL 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 65 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	63.5	19	4	US-09-959-873B-8
2	101	63.5	34	4	US-09-959-873B-9
3	101	63.5	54	4	US-09-438-833-13
4	101	63.5	116	4	US-09-438-833-8
5	101	63.5	288	4	US-09-438-833-9
6	101	63.5	301	4	US-09-438-833-10
7	101	63.5	311	4	US-09-438-833-7
8	101	63.5	532	4	US-09-949-016-7389
9	101	63.5	613	4	US-09-438-833-6
10	101	63.5	652	4	US-09-438-833-5
11	101	63.5	756	4	US-09-438-833-11
12	101	63.5	805	2	US-08-480-473B-4
13	101	63.5	805	3	US-08-915-213-4
14	101	63.5	805	3	US-09-235-217-4
15	101	63.5	805	5	PCT-US96-10251-4
16	101	63.5	810	1	US-08-785-241-7
17	101	63.5	813	4	US-09-438-833-12
18	101	63.5	826	1	US-08-785-241-6
19	101	63.5	826	2	US-08-480-473B-2
20	101	63.5	826	3	US-08-915-213-2
21	101	63.5	826	3	US-09-148-547-2
22	101	63.5	826	3	US-09-235-217-2
23	101	63.5	826	3	US-09-380-662-23
24	101	63.5	826	4	US-09-438-833-1
25	101	63.5	826	4	US-09-702-703-330
26	101	63.5	826	4	US-09-736-457-330
27	101	63.5	826	4	US-09-383-581-2

28	101	63.5	826	4	US-09-614-124B-330	Sequence 330, App
29	101	63.5	826	4	US-09-671-325-330	Sequence 330, App
30	101	63.5	826	4	US-09-589-184-330	Sequence 330, App
31	101	63.5	826	4	US-09-658-824-330	Sequence 330, App
32	101	63.5	826	4	US-09-959-873B-18	Sequence 18, Appl
33	101	63.5	826	4	US-09-949-016-6089	Sequence 6089, Ap
34	101	63.5	826	4	US-09-967-388-4	Sequence 4, Appl
35	101	63.5	826	5	PCT-US96-10251-2	Sequence 2, Appl
36	101	63.5	827	4	US-09-919-039-149	Sequence 149, App
37	89	56.0	19	4	US-09-972-784-5	Sequence 5, Appl
38	76.5	48.1	205	3	US-09-374-454-2	Sequence 2, Appl
39	76.5	48.1	870	1	US-08-785-241-4	Sequence 4, Appl
40	76.5	48.1	870	3	US-09-374-454-6	Sequence 6, Appl
41	76.5	48.1	875	1	US-08-785-241-5	Sequence 5, Appl
42	75	47.2	385	1	US-08-450-257-58	Sequence 58, Appl
43	75	47.2	385	1	US-08-450-246-58	Sequence 58, Appl
44	75	47.2	385	1	US-08-450-098-58	Sequence 58, Appl
45	75	47.2	385	1	US-08-451-233-58	Sequence 58, Appl
46	75	47.2	385	1	US-08-450-236-58	Sequence 58, Appl
47	75	47.2	385	3	US-08-235-403-58	Sequence 58, Appl
48	60	37.7	32	2	US-08-706-741B-87	Sequence 87, Appl
49	60	37.7	32	2	US-08-924-695A-87	Sequence 87, Appl
50	60	37.7	134	1	US-08-450-257-38	Sequence 38, Appl
51	60	37.7	134	1	US-08-450-246-38	Sequence 38, Appl
52	60	37.7	134	1	US-08-450-098-38	Sequence 38, Appl
53	60	37.7	134	1	US-08-451-233-38	Sequence 38, Appl
54	60	37.7	134	1	US-08-450-236-38	Sequence 38, Appl
55	60	37.7	134	3	US-08-235-403-38	Sequence 38, Appl
56	60	37.7	143	1	US-08-450-257-63	Sequence 63, Appl
57	60	37.7	143	1	US-08-450-246-63	Sequence 63, Appl
58	60	37.7	143	1	US-08-450-098-63	Sequence 63, Appl
59	60	37.7	143	1	US-08-451-233-63	Sequence 63, Appl
60	60	37.7	143	1	US-08-450-236-63	Sequence 63, Appl
61	60	37.7	143	3	US-08-235-403-63	Sequence 63, Appl
62	59	37.1	28	3	US-09-041-886-50	Sequence 50, Appl
63	59	37.1	28	3	US-09-041-886-56	Sequence 56, Appl
64	58	36.5	11	2	US-08-706-741B-54	Sequence 54, Appl
65	58	36.5	11	2	US-08-924-695A-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-09-959-873B-8
; Sequence 8, Application US/09959873B
; Patent No. 6787226
; GENERAL INFORMATION:
; APPLICANT: Ratcliffe, Peter John
; APPLICANT: Maxwell, Patrick Henry
; APPLICANT: Pugh, Christopher William
; TITLE OF INVENTION: Interaction Between the VHL Tumour
; TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods
; TITLE OF INVENTION: Relating Thereto
; FILE REFERENCE: 3547.1000-000
; CURRENT APPLICATION NUMBER: US/09/959,873B
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: PCT/GB00/01826
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: GB9911047.0
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Motif
US-09-959-873B-8
Query Match 63.5%; Score 101; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;

```
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 1 DLDLEMLAPYIPMDDDFQL 19

RESULT 2
US-09-959-873B-9
; Sequence 9, Application US/09959873B
; Patent No. 6787326
; GENERAL INFORMATION:
; APPLICANT: Ratcliffe, Peter John
; APPLICANT: Maxwell, Patrick Henry
; APPLICANT: Pugh, Christopher William
; TITLE OF INVENTION: Interaction Between the VHL Tumour
; TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods
; TITLE OF INVENTION: Relating Thereto
; FILE REFERENCE: 3547.1000-000
; CURRENT APPLICATION NUMBER: US/09/959,873B
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: PCT/GB00/01826
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: GB9911047.0
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-959-873B-9

Query Match 63.5%; Score 101; DB 4; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 8 DLDLEMLAPYIPMDDDFQL 26

RESULT 3
US-09-438-833-13
; Sequence 13, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
US-09-438-833-13

Query Match 63.5%; Score 101; DB 4; Length 54;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 26 DLDLEMLAPYIPMDDDFQL 44
```

```
RESULT 4
US-09-438-833-8
; Sequence 8, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
US-09-438-833-8

Query Match 63.5%; Score 101; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 31 DLDLEMLAPYIPMDDDFQL 49

RESULT 5
US-09-438-833-9
; Sequence 9, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
US-09-438-833-9

Query Match 63.5%; Score 101; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 31 DLDLEMLAPYIPMDDDFQL 49

RESULT 6
US-09-438-833-10
; Sequence 10, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
US-09-438-833-10
```

```

; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 5
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 1-652 of human HIF-1 alpha
US-09-438-833-5

      Query Match          63.5%; Score 101; DB 4; Length 652;
      Best Local Similarity 100.0%; Pred. No. 9.2e-07;
      Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12  DLDEMLAPYIPMDDDFQL 30
          |||||
          556  DLDEMLAPYIPMDDDFQL 574

Db

RESULT 11
US-09-438-833-11
; Sequence 11, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants

```

us-10-032-361-7.rai

Wed Feb 9 06:58:02 2005

```

; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 71-826 of human HIF-1 alpha
US-09-438-833-11

Query Match      63.5%; Score 101; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12 DLDLEMLAPYIPMDDDFQL 30
      |||||
Db      486 DLDLEMLAPYIPMDDDFQL 504

RESULT 12
US-08-480-473B-4
; Sequence 4, Application US/08480473B
; Patent No. 5882914
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,473B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-473B-4

Query Match      63.5%; Score 101; DB 2; Length 805;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12 DLDLEMLAPYIPMDDDFQL 30
      |||||
Db      535 DLDLEMLAPYIPMDDDFQL 553

RESULT 13
US-08-480-473B-4
; Sequence 4, Application US/08480473B
; Patent No. 5882914
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/235,217
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-473B-4

Query Match      63.5%; Score 101; DB 3; Length 805;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12 DLDLEMLAPYIPMDDDFQL 30
      |||||
Db      535 DLDLEMLAPYIPMDDDFQL 553

RESULT 14
US-09-235-217-4
; Sequence 4, Application US/09235217
; Patent No. 6222018
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/235,217
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-235-217-4

```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-235-217-4

Query Match 63.5%; Score 101; DB 3; Length 805;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DLDLEMLAPYIPMDDDFQL 30
DB 535 DLDLEMLAPYIPMDDDFQL 553

RESULT 15
PCT-US96-10251-4
; Sequence 4, Application PC/TUS9610251
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10251
; FILING DATE: 06-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5099
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-10251-4

Query Match 63.5%; Score 101; DB 5; Length 805;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-235-217-4
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-235-217-4

Query Match 63.5%; Score 101; DB 3; Length 805;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DLDLEMLAPYIPMDDDFQL 30
DB 535 DLDLEMLAPYIPMDDDFQL 553

RESULT 16
US-08-785-241-7
; Sequence 7, Application US/08785241
; Patent No. 5695963
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; TITLE OF INVENTION: Endothelial PAS Domain Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,241
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1229
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 810 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-785-241-7

Query Match 63.5%; Score 101; DB 1; Length 810;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DLDLEMLAPYIPMDDDFQL 30
DB 543 DLDLEMLAPYIPMDDDFQL 561

RESULT 17
US-09-438-833-12
; Sequence 12, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

us-10-032-361-7-rai

Wed Feb 9 06:58:02 2005

OTHER INFORMATION: Description of Artificial Sequence: Subdomain
; OTHER INFORMATION: 1-813 of human HIF-1 alpha
US-09-438-833-12

Query Match 63.5%; Score 101; DB 4; Length 813;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 18
US-08-785-241-6
; Sequence 6, Application US/08785241
; Patent No. 5695963
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; APPLICANT: Tian, Hui
; TITLE OF INVENTION: Endothelial PAS Domain Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,241
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UTSD:1229
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-785-241-6

Query Match 63.5%; Score 101; DB 1; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 19
US-08-480-473B-2
; Sequence 2, Application US/08480473B
; Patent No. 5882914
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:

ADDRESS: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,473B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/053001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 826 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-473B-2

Query Match 63.5%; Score 101; DB 2; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 20
US-08-915-213-2
; Sequence 2, Application US/08915213
; Patent No. 6020462
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,213
FILING DATE: 20-AUG-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,473
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/053001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 826 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-915-213-2
;
; Query Match      63.5%; Score 101; DB 3; Length 826;
; Best Local Similarity 100.0%; Pred. No. 1.2e-06;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 21
US-09-148-547-2
; Sequence 2, Application US/09148547
; Patent No. 6124131
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: Hypoxia Inducible Factor-1 and Methods of Use
; FILE REFERENCE: 07265/151001
; CURRENT APPLICATION NUMBER: US/09/148,547
; CURRENT FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-148-547-2

Query Match      63.5%; Score 101; DB 3; Length 826;
; Best Local Similarity 100.0%; Pred. No. 1.2e-06;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 22
US-09-235-217-2
; Sequence 2, Application US/09235217
; Patent No. 622018
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/235,217
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,473
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.

; INFORMATION FOR SEQ ID NO: 2:
; REFERENCE/DOCKET NUMBER: 07265/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 826 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-235-217-2
;
; Query Match      63.5%; Score 101; DB 3; Length 826;
; Best Local Similarity 100.0%; Pred. No. 1.2e-06;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 23
US-09-380-662-23
; Sequence 23, Application US/09380662
; Patent No. 6376199
; GENERAL INFORMATION:
; APPLICANT: Caniggia, Isabella
; APPLICANT: Post, Martin
; APPLICANT: Lye, Stephen
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF TROPHOBLAST
; FILE REFERENCE: 11757.38USWO
; CURRENT APPLICATION NUMBER: US/09/380,662
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-380-662-23

Query Match      63.5%; Score 101; DB 3; Length 826;
; Best Local Similarity 100.0%; Pred. No. 1.2e-06;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 24
US-09-438-833-1
; Sequence 1, Application US/09438833
; Patent No. 6436654
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; TITLE OF INVENTION: Protein variants
; FILE REFERENCE: 1848
; CURRENT APPLICATION NUMBER: US/09/438,833
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
```

; VOLUME: 92
; PAGES: 5510-5514
; DATABASE ACCESSION NUMBER: GenBank U22431
; DATABASE ENTRY DATE: 1995-06-28
US-09-438-833-1

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 25
US-09-702-705-330
; Sequence 330, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-705-330

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 26
US-09-736-457-330
; Sequence 330, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-330

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 27
US-09-383-581-2
; Sequence 2, Application US/09383581
; Patent No. 6562799
; GENERAL INFORMATION:
; APPLICANT: Semenza, Gregg L.
; TITLE OF INVENTION: STABLE HYPOXIA INDUCIBLE FACTOR-1 alpha
; TITLE OF INVENTION: AND METHOD OF USE
; FILE REFERENCE: JHUI500-1
; CURRENT APPLICATION NUMBER: US/09/383,581
; CURRENT FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/148,547
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-383-581-2

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 28
US-09-614-124B-330
; Sequence 330, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-330

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;


```
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 29
US-09-671-325-330
; Sequence 330, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-325-330

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 30
US-09-589-184-330
; Sequence 330, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-589-184-330

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 31
US-09-658-824-330
; Sequence 330, Application US/09658824
; Patent No. 6746846
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C11
; CURRENT APPLICATION NUMBER: US/09/658,824
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 1788
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-658-824-330

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 32
US-09-959-873B-18
; Sequence 18, Application US/09959873B
; Patent No. 6787326
; GENERAL INFORMATION:
; APPLICANT: Ratcliffe, Peter John
; APPLICANT: Maxwell, Patrick Henry
; APPLICANT: Pugh, Christopher William
; TITLE OF INVENTION: Interaction Between the VHL Tumour
; TITLE OF INVENTION: Suppressor and Hypoxia Inducible Factor, and Assay Methods
; TITLE OF INVENTION: Relating Thereto
; FILE REFERENCE: 3547.1000-000
; CURRENT APPLICATION NUMBER: US/09/959,873B
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: PCT/GB00/01826
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: GB9911047.0
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-959-873B-18

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
Db 556 DLDLEMLAPYIPMDDDFQL 574
```

```
Db 556 DLDLEMLAPYIPMDDDFQL 574
;
; ADDRESSER: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10251
; FILING DATE: 06-JUN-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/053WO1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-10251-2

Query Match 63.5%; Score 101; DB 5; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
| | | | | | | | | | | | | | | |
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 36
US-09-919-039-149
; Sequence 149, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 1250434CD1
US-09-919-039-149

Query Match 63.5%; Score 101; DB 4; Length 827;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
| | | | | | | | | | | | | | | |
Db 557 DLDLEMLAPYIPMDDDFQL 575

RESULT 37
US-09-972-784-5
```

```
Db 556 DLDLEMLAPYIPMDDDFQL 574
;
; Sequence 6089, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6089
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-6089

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
| | | | | | | | | | | | | | | |
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 34
US-09-967-388-4
; Sequence 4, Application US/09967388
; Patent No. 6838430
; GENERAL INFORMATION:
; APPLICANT: JEFFEREY M. ARBEIT
; TITLE OF INVENTION: USE OF HIF-1ALPHA VARIANTS TO ACCELERATE
; WOUND HEALING
; FILE REFERENCE: UC077.001A
; CURRENT APPLICATION NUMBER: US/09/967,388
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HUMAN
; US-09-967-388-4

Query Match 63.5%; Score 101; DB 4; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30
| | | | | | | | | | | | | | | |
Db 556 DLDLEMLAPYIPMDDDFQL 574

RESULT 35
PCT-US96-10251-2
; Sequence 2, Application PC/TUS9610251
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
```

; Sequence 5, Application US/09972784

; Patent No. 6586088

; GENERAL INFORMATION:

; APPLICANT: McKnight, Steven L.

; APPLICANT: Bruik, Richard K.

; TITLE OF INVENTION: Prolyl-4-Hydroxylases

; FILE REFERENCE: UTS0871

; CURRENT APPLICATION NUMBER: US/09/972,784

; CURRENT FILING DATE: 2001-10-04

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 19

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: sequence derived from HIF-a ODD domain.

US-09-972-784-5

Query Match

Best Local Similarity 56.0%; Score 89; DB 4; Length 19;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 12 DLDLEMLAPYIPMDDDFQL 30

Db 1 DLDLEALAPYIPADDDFQL 19

RESULT 38

US-09-374-454-2

; Sequence 2, Application US/09374454

; Patent No. 6395548

; GENERAL INFORMATION:

; APPLICANT: Lee, Mu-En

; APPLICANT: Maemura, Koji

; APPLICANT: Hsieh, Chung-Ming

; TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS

; FILE REFERENCE: 05433/037001

; CURRENT APPLICATION NUMBER: US/09/374,454

; CURRENT FILING DATE: 1999-08-13

; EARLIER APPLICATION NUMBER: US 60/096,515

; EARLIER FILING DATE: 1998-08-14

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 205

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-374-454-2

Query Match

Best Local Similarity 48.1%; Score 76.5; DB 3; Length 205;

Matches 16; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 12 DLDLEMLAPYIPMD-DDFQL 30

Db 38 ELDLETLAPYIPMDGEDFQL 57

RESULT 39

US-08-785-241-4

; Sequence 4, Application US/08785241

; Patent No. 5695963

; GENERAL INFORMATION:

; APPLICANT: McKnight, Steven L.

; APPLICANT: Russell, David W.

; APPLICANT: Tian, Hui

; TITLE OF INVENTION: Endothelial PAS Domain Protein

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/785,241

; FILING DATE: 17-JAN-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UTS01229

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 870 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-785-241-4

Query Match

Best Local Similarity 48.1%; Score 76.5; DB 1; Length 870;

Matches 16; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 12 DLDLEMLAPYIPMD-DDFQL 30

Db 523 ELDLETLAPYIPMDGEDFQL 542

RESULT 40

US-09-374-454-6

; Sequence 6, Application US/09374454

; Patent No. 6395548

; GENERAL INFORMATION:

; APPLICANT: Lee, Mu-En

; APPLICANT: Maemura, Koji

; APPLICANT: Hsieh, Chung-Ming

; TITLE OF INVENTION: METHODS OF MODULATING OF ANGIOGENESIS

; FILE REFERENCE: 05433/037001

; CURRENT APPLICATION NUMBER: US/09/374,454

; CURRENT FILING DATE: 1999-08-13

; EARLIER APPLICATION NUMBER: US 60/096,515

; EARLIER FILING DATE: 1998-08-14

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 870

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-374-454-6

Query Match

Best Local Similarity 48.1%; Score 76.5; DB 3; Length 870;

Matches 16; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 12 DLDLEMLAPYIPMD-DDFQL 30

Db 523 ELDLETLAPYIPMDGEDFQL 542

RESULT 41

US-08-785-241-5

; Sequence 5, Application US/08785241

; Patent No. 5695963

GENERAL INFORMATION:
APPLICANT: McKnight, Steven L.
APPLICANT: Russell, David W.
APPLICANT: Tian, Hui
TITLE OF INVENTION: Endothelial PAS Domain Protein
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,241
FILING DATE: 17-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UTSD:1229
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 875 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-785-241-5

Query Match 48.1%; Score 76.5; DB 1; Length 875;
Best Local Similarity 80.0%; Pred. No. 0.0058;
Matches 16; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 12 DLDLEMLAPYIPMD-DDFOL 30
Db 523 ELDLETLAPYIPMDGEDFOL 542

RESULT 42
US-08-450-257-58
Sequence 58, Application US/08450257
Patent No. 5652122
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: PABO, Carl
APPLICANT: BARSOUM, James G.
APPLICANT: FAWELL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,257

FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454,450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636,662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158,015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-450-257-58

Query Match 47.2%; Score 75; DB 1; Length 385;
Best Local Similarity 66.7%; Pred. No. 0.0038;
Matches 16; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRRLDLEMLAPYIPM 24
Db 2 YGRKKRQRRRLPSQAQLMFPSPM 25

RESULT 43
US-08-450-246-58
Sequence 58, Application US/08450246
Patent No. 5670617
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: PABO, Carl
APPLICANT: BARSOUM, James G.
APPLICANT: FAWELL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,246
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403